

Residue Identity = 100%	Matches = 349	Mismatches = 0
Gaps = 0	Conservative Substitutions = 0	
X 10 20 30 40 50 60 70		
MTAASMGVRVAFVVLALCSRPAGVQNCSPCRCPDEPAPRCFAGVSLVLDGGCCRCVCAKQGLCTERD		
MTAASMGVRVAFVVLALCSRPAGVQNCSPCRCPDEPAPRCFAGVSLVLDGGCCRCVCAKQGLCTERD		
X 10 20 30 40 50 60 70		
PCDPHKGLFCDFGSPANKIGVCTAKGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLP		
PCDPHKGLFCDFGSPANKIGVCTAKGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLP		
80 90 100 110 120 130 140		
SPDCPFPRRVKLPKGCCEEWVCDPEPKQTVVGPALAAAYRLDTFGPDPTMIRANCLVQTTEWSACSKTCGMG		
SPDCPFPRRVKLPKGCCEEWVCDPEPKQTVVGPALAAAYRLDTFGPDPTMIRANCLVQTTEWSACSKTCGMG		
150 160 170 180 190 200 210		
ISTRVTNDNASCRLEKOSRLCMWRPCEADLEENIKKGGKCIPTPKISKPIKFKELSGCTSMKTYRAKFCGVCT		
ISTRVTNDNASCRLEKOSRLCMWRPCEADLEENIKKGGKCIPTPKISKPIKFKELSGCTSMKTYRAKFCGVCT		
220 230 240 250 260 270 280		
DGRCTPHRTTTLPEVFKCPDGEVKNKMMFIKTCACHYNCPCGNDNIFESLYYRMYGDMA		
DGRCTPHRTTTLPEVFKCPDGEVKNKMMFIKTCACHYNCPCGNDNIFESLYYRMYGDMA		
290 300 310 320 330 340 X		

Specter 386680

TAYLOR
***** (TM)

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Sep 17 09:35:08 1997; MasPar time 7.87 Seconds
507.112 Million cell updates/sec
Tabular output not generated.

Title: >US-08-167-628-2
Description: (1-349) from 3408040.pep
Perfect Score: 2713
Sequence: 1 MTASMGPRVAFVLLALC.....PGNDIFESLYRKMVGMDA 349

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.775; Variance 141.735; scale 0.231
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Pred. No.
1	2713	100.0	349	20 W11302	Connective tissue gro 3.93e-253
2	2713	100.0	349	16 R79964	Connective tissue gro 3.93e-253
3	2521	92.9	348	5 R25566	Beta-IG-M2. 4.65e-234
4	1460	53.8	351	6 R31599	Chicken nov protein. 4.62e-129
5	1113	41.0	379	5 R25565	Beta-IG-M1. 5.50e-95
6	1002	36.9	375	16 R90919	Connective tissue gro 3.88e-84
7	644	23.7	205	6 R31608	Homologous to chicken 1.93e-49

8	579	21.3	72	6	R31610	Fragment XXVI with ho
9	430	15.8	84	6	R31602	Chicken nov protein f
10	412	15.2	71	6	R31609	Encoded by chicken no
11	405	14.9	84	6	R31603	Polyptide X homolog
12	395	14.6	124	10	R46078	CYR61 like protein.
13	314	11.6	72	6	R31605	Homologous to chicken
14	312	11.5	70	6	R31604	Chicken nov protein f
15	278	10.2	75	6	R31601	Chicken nov protein f
16	268	9.9	76	6	R31600	Chicken nov protein f
17	157	5.8	22	6	R31612	Fragment XXXI homolog
18	138	5.1	22	6	R31611	Fragment XXX encoded
19	125	4.6	271	5	R26994	Rat IGFBP-5.
20	123	4.5	184	19	R98994	Vascular IBP-like gro
21	123	4.5	238	4	R22533	Sequence of insulin-1
22	123	4.5	238	4	R21688	Sequence of insulin-1
23	121	4.5	455	19	W00231	Drosophila morphogeni
24	121	4.5	455	9	R47261	Pre-pro 60A.
25	121	4.5	455	10	R60966	Drosophila 60A morpho
26	121	4.5	455	10	R37981	Drosophila 60A morpho
27	121	4.5	455	7	R33413	Drosophila 60A morpho
28	121	4.5	455	7	R33935	Drosophila 60A full le
29	121	4.5	455	9	R47298	Morphogen 60A full le
30	121	4.5	455	10	R50208	Drosophila 60A morpho
31	121	4.5	455	10	R60962	Drosophila 60A morpho
32	121	4.5	455	9	R46745	Drosophila sp. 60(A)
33	117	4.3	282	13	R79101	Prostaglandin I2 (PGI
34	117	4.3	282	13	R79102	Prostaglandin I2 (PGI
35	113	4.2	193	3	P60463	Sequence of C-terminu
36	113	4.2	2813	3	P60462	Sequence of human von
37	112	4.1	2813	3	P60053	Sequence of von Wille
38	109	4.0	1218	19	W05833	Human Serrate-1 (HJ1)
39	105	3.9	272	19	R95329	Insulin-like growth f
40	105	3.9	272	5	R26995	Human IGFBP-5.
41	105	3.9	272	5	R25700	IGFBP6.
42	105	3.9	272	10	R55084	Human insulin-like gr
43	106	3.9	807	8	R44241	F-spondin (FP5-9).
44	103	3.8	328	1	R04908	EcoRI-EcoRI fragment
45	103	3.8	532	17	R83016	Recombinant papilloma

ALIGNMENTS

RESULT	1
ID	W11302 standard; Protein; 349 AA.
AC	W11302;
DT	18-MAR-1997 (first entry)
DE	Connective tissue growth factor.
KW	Connective tissue growth factor; CTGF; human; connective tissue cell;
KW	proliferative disease; platelet-derived growth factor; PDGF; development;
KW	tissue growth; repair; umbilical vein endothelial cell; HUVE cell;
KW	antibody; wound healing; cancer; fibrotic disease; atherosclerosis;
KW	inhibitor; protease degradation; growth factor; therapy.
OS	Homo sapiens.
PN	US5393240-A.
PD	1996-08-14
PF	30-AUG-1991; 752427.
PR	30-AUG-1991; US-752427.
PR	14-DEC-1993; US-167628.
PR	10-FEB-1995; US-386680.
PA	(OYSP-) UNIV SOUTH FLORIDA.

PI Bradham DM, Grotendorst GR;
DR WPI; 97-051180/05.
PT New nucleic acid encoding connective tissue growth factor - useful
PT for accelerating wound healing, also for diagnosis and treatment of
PT proliferative disease
PS Claim 9; Column 15-18; 11pp; English.
CC This sequence represents the human connective tissue growth factor
CC (CTGF). CTGF is related immunologically and biologically to
CC platelet-derived growth factor (PDGF), but is encoded by an unrelated
CC gene. CTGF is thought to play a significant role in the normal
CC development, growth, and repair of human tissue, similarly to PDGF. The
CC cDNA encoding this sequence was isolated by screening a cDNA library from
CC human umbilical vein endothelial (HUVE) cells with anti-PDGF antibodies.
CC CTGF may be used to accelerate wound healing. Also, elevated levels of
CC CTGF may be diagnostic of proliferative diseases involving outgrowth of
CC connective tissue cells, such as cancer, fibrotic disease and
CC atherosclerosis. All of these diseases can be treated with reagents
CC reactive with CTGF, such as antibodies (which can also serve as assay
CC reagents). Antisense nucleic acids, and ribozymes could also be used to
CC inhibit CTGF production. The advantage with using CTGF is that it is
CC more stable, and less susceptible to protease degradation than PDGF, and
CC other growth factors involved in wound healing. This is believed to be
CC due to the high Cys content.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 20; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps D;
Db 1 mtaasngpvrvaavfvlalcsrpavqncsgpcrcpdeprapcpagvslvdgcgcrcvc 60
Qy 1 MTAASMGPPVRVAFVWVLLALCSRPVAVQNCSCGPCRCPDEPRAPCPAGVSLVDGCGCCRCV 60
Db 61 akqlgelcterdpcdpkhgldfcdgspanrkigvctakdgapciffgtvyrsgesfqssc 120
Qy 61 AKQLGELCTERDPCDPKHGLFCDFGSPANRKIGVCTAKDGAPCIFGTVYRSGESFQSSC 120
Db 121 kyqctcldgavgcmlcsmdvrlpsdpcpfprvrklpgkceewvcdpdkdtvvgpala 180
Qy 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCPFPRVRKLPKGCCEEWVCDPDKDTVVGPA 180
Db 181 ayrledtfgpdpmtiranclvqtewsacsktcgmigistrvndnasrlekqslcmvr 240
Qy 181 AYRLDTFGPDPTMIRANCLVQTTEWSACSKTCGMISTRVNDNASCRLEKQSLCMVR 240
Db 241 pceadleenikkkgkcirtpkiskpikfelsgctsmktyrakfcgvctdgrcctphrttt 300
Qy 241 PCEADLEENIKKGKCIKIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCTPHRTTT 300
Db 301 lpvefkcpdgvemkmmfiktachyncpgndndifeslyyrkmygdma 349
Qy 301 LPVEFKCPDGEVEMKMMFIKTCACHYNCPCNDNDFESLYYRMYGDMA 349

RESULT 2
ID R79964 standard; Protein; 349 AA.
AC R79964;
DT 12-JUN-1996 (first entry)
DE Connective tissue growth factor.

KW Connective tissue growth factor; CTGF; wound healing; vulnary;
KW cell proliferation; cancer; fibrosis; atherosclerosis; diagnosis
OS therapy; mitogen.
FH Homo sapiens.

Key Location/Qualifiers
FT Modified site 28
FT /label= N-glycosylation_site
FT Modified site 225
FT /label= N-glycosylation_site
PN US408040-A.
PD 18-APR-1995.
PF 30-AUG-1991; 752427.
PR 30-AUG-1991; US-752427.
PR 14-DEC-1993; US-167628.
PA (YISF-) UNIV SOUTH FLORIDA.
PI Bradham DM, Grotendorst GR;
DR WPI; 95-161147/21.
DR N-PSDB; T04226.
PT New connective tissue growth factor - used to develop prods. for
PT wound healing and for diagnosis and therapy of cell proliferative
PT disorders.
PS Claim 1; Column 19-20; 12pp; English.
CC Novel human connective tissue growth factor (CTGF) (R79964)
CC is related immunologically and biologically to platelet-derived
CC growth factor (PDGF), but is the product of a distinct gene.
CC CTGF is mitogenic and also a chemotactic agent for cells. It is
CC produced by endothelial and fibroblastic cells, and probably acts
CC as a growth factor in wound healing. Recombinant CTGF can be obtd.
CC by expression of cDNA clone DB60R32 (T04226) in transformed host
CC cells. It is used to accelerate wound healing, and to raise
CC antibodies useful in detecting disorders associated with overgrowth
CC of cells, such as cancer, fibrotic diseases and atherosclerosis.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 16; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasngpvrvaavfvlalcsrpavqncsgpcrcpdeprapcpagvslvdgcgcrcvc 60
Qy 1 MTAASMGPPVRVAFVWVLLALCSRPVAVQNCSCGPCRCPDEPRAPCPAGVSLVDGCGCCRCV 60
Db 61 akqlgelcterdpcdpkhgldfcdgspanrkigvctakdgapciffgtvyrsgesfqssc 120
Qy 61 AKQLGELCTERDPCDPKHGLFCDFGSPANRKIGVCTAKDGAPCIFGTVYRSGESFQSSC 120
Db 121 kyqctcldgavgcmlcsmdvrlpsdpcpfprvrklpgkceewvcdpdkdtvvgpala 180
Qy 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCPFPRVRKLPKGCCEEWVCDPDKDTVVGPA 180
Db 181 ayrledtfgpdpmtiranclvqtewsacsktcgmigistrvndnasrlekqslcmvr 240
Qy 181 AYRLDTFGPDPTMIRANCLVQTTEWSACSKTCGMISTRVNDNASCRLEKQSLCMVR 240
Db 241 pceadleenikkkgkcirtpkiskpikfelsgctsmktyrakfcgvctdgrcctphrttt 300
Qy 241 PCEADLEENIKKGKCIKIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCTPHRTTT 300
Db 301 lpvefkcpdgvemkmmfiktachyncpgndndifeslyyrkmygdma 349
Qy 301 LPVEFKCPDGEVEMKMMFIKTCACHYNCPCNDNDFESLYYRMYGDMA 349

QY 301 LPVEFKCPDGEVVKNNMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349

RESULT 3

ID R25566 standard; Protein; 348 AA.
 AC R25566;
 DT 18-JAN-1993 (first entry)
 DE Beta-IG-M2.
 KW Transforming growth factor beta; induced; CEF-10; v-src; chicken;
 KW embryo; fibroblasts; TGF-beta.
 OS Mus musculus.
 PN EP-495674-A.
 PD 22-JUL-1992.
 PF 17-JAN-1992; 300429.
 PR 18-JAN-1991; US-842991.
 PR 10-JAN-1992; US-816270.
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
 DR WEI; 92-243508/30.
 DR N-PSDB; Q26422.
 PT TGF-beta induced gene family - encodes proteins involved in
 PT growth and differentiation effects of TGF-beta-1
 PS Claim 3; Fig 2; 35pp; English.
 CC The protein sequence was deduced from the DNA sequence obtd. by
 CC screening a cDNA library made from AKR-2B mouse cells induced with
 CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
 CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
 CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
 CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
 CC Beta-IG-M2 displays 50 percent homology to the CEF-10 protein
 CC induced by v-src in chicken embryo fibroblasts. Residues 52-59
 CC of beta-IG-M2 conform to the GGGCCXC motif reported in the
 CC amino half of insulin-like growth factor (IGF) binding proteins.
 CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
 CC an amino acid sequence with strong homology to a motif found near the
 CC C-terminal of the malarial circumsporozoite (CS) protein, which is
 CC highly conserved among all species of malarial parasites sequenced
 CC to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate
 CC cell-cell and cell-extracellular matrix interactions, such as
 CC properdin, thrombospondin, and TRAP. The proteins encoded by
 CC TGF-beta induced genes are likely to be involved in mediation of
 CC the biological effects of TGF-beta relating to cell growth and
 CC differentiation. See also R25565.
 SQ Sequence 348 AA;

Query Match 92.9%; Score 2521; DB 5; Length 348;
 Best Local Similarity 91.1%; Pred. No. 4.65e-234;
 Matches 316; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

Db 1 mlavagplslalv-llalctrptgqdsacqcaaeaphcpagvslvldgcgcrcv 59
 QY 1 MTAASMGPRVAFVLLALCSRPVGNQSCGPCRDEPAPCPAGVSLVLDGCGCRVC 60
 Db 60 aklgelcterdpchphkglfdcgspnarkigvctakdgapcvfgsvrgesfgssc 119
 QY 61 AKQGLGELCTERDPCDPRHGLFCDFGSPANRKIGVCTARDGAPCIFGTVYRSGESFQSSC 120
 Db 120 kyqctcldgavcgvplcamdvrlbpdcpfrrvklpgkceewvcddepkdrtaavgpala 179
 |||||

QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSPDCPFRVRVKLPKGCCEBWCDPEKQTVVGPALA 180
 Db 180 ayrlctfsgdpdptmmrancelvqtewsacaktcmgmistrtndntcfclqksrlcmvr 239
 QY 181 AYRLCTFSGDPDPTMIRANCLVQVTEWSACSKTCGMGISTRTVTDNASCRLKQSRLCMYR 240
 Db 240 pceadleenikkgkckirtkpkvkvkfslgctsvktyrakfcgvtcdgrctcphrttt 299
 QY 241 PCEADLEENIKKGKCKIRTPKIRKIFELSGCTSMKTYRAKFCGVTCDGRCTCPHRTTT 300
 Db 300 lpvefkcpdgeimkkmnmfiktcachnycpgdndifeslyyrmkmygdma 348
 QY 301 LPVEFKCPDGEVVKNNMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349

RESULT 4

ID R31599 standard; Protein; 351 AA.
 AC R31599;
 DT 24-MAY-1993 (first entry)
 DE Chicken nov protein.
 KW avian nephroblastoma; avian myeloblastoma virus; IGF binding site;
 KW insulin-like growth factor; Wilm's tumour.
 OS Gallus domesticus.
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal peptide
 FT /note= "only hydrophobic region of protein"
 FT Binding site 56..63
 FT /label= IGF-binding site motif
 FT /note= "corresponds to GGGCCXC consensus"
 FN W09300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C, Perbal B;
 DR WFI; 93-036377/04.
 DR N-PSDB; Q36031.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 1; Fig 1; 67pp; French.
 CC This amino acid sequence was deduced from the nucleotide sequence
 CC of a chicken nov gene clone isolated from a gene bank prepared from
 CC chicken embryonic fibroblasts screened with a tumour-derived probe.
 CC The only hydrophobic region occurs within the putative signal
 CC peptide suggesting that the protein is secreted. The protein also
 CC contains the consensus motif of proteins which bind to insulin-like
 CC growth factors. It is known that the human IGFII gene is
 CC overexpressed in some Wilm's tumours and a similar deregulation of
 CC IGFII expression could be involved in nephroblastoma development.
 CC The deduced nov protein sequence contains 39 (non-clustered)
 CC cysteine residues.
 SQ Sequence 351 AA;

Query Match 53.8%; Score 1460; DB 6; Length 351;
 Best Local Similarity 54.2%; Pred. No. 4.62e-129;
 Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;

Db 3 tgggggplvllllllllrrpcevsgreaacprcpggrcraep-prcapgvavldgcgccl 61

```

QY 2 TAASMG-PVRVAFVLLALCSREAVGQNCSGFC--RCDEPAPRCAGVSLVLDGCGCR 58
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 62 veargrescpplpcedgglycdrgpddggagimvlegdhcnvfdgmlyrnetgfcqp 121
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 59 VCAKQLGELCTERDPCDPKHGLFCDFGSPANRKGIVGTAKDGAPCFGFTYRSGESFQS 118
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 122 sckyqctordgigelpcnlglilpgpdcfpkrievpgceckwcd-prdevlllgff 180
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 119 SKKYQCTCLDGAAGVCMPLCSMDVRLPSPDCPPFRVRKLPKGCCEEWCDKPDQTVVGP- 177
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 181 amaavqeatglidvdsaaancleqtewsacskcmgfstvtrnnrqcmvktqlrc 240
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 178 ALAAAYLEDTFDPDMIRANCLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRLC 237
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 241 mmrprcene-epedkkgkciqtckemkavrfeykncvavtqykypryogclndgrcctphn 299
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 238 MVRPCEADLEENIKKGGKIRTPKISKP IKFELSGCTSMKTVRAKFCGVTGRCCTPHR 297
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 30D tktiqlvfcpggkflkpmmlintevchngncpqnannaffqpl 342
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 298 TTTLPVEFKPDGEVMMKNMFIKTCACHYNCP-GDNDFESL 339
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |

```

RESULT 5

```

ID R25565 standard; Protein; 379 AA.
AC R25565;
DT 18-JAN-1993 (first entry)
DE Beta-IG-M1.
KW Transforming growth factor beta; induced; CEF-10; v-src; chicken;
KW embryo; fibroblasts; TGF-beta.
OS Mus musculus.
PN EP-495674-A.
PD 22-JUL-1992.
PR 17-JAN-1991; US-642991.
PR 18-JAN-1991; US-616270.
PR 10-JAN-1992; US-816270.
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
DR WPI; 92-243308/30.
DR N-PSDB; Q26421.
PT TGF-beta induced gene family - encodes proteins involved in
PT growth and differentiation effects of TGF-beta-1
PS Claim 2; Fig 1; 3pp; English.
CC The protein sequence was deduced from the DNA sequence obt'd. by
CC screening a cDNA library made from AKR-2B mouse cells induced with
CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
CC Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
CC induced by v-src in chicken embryo fibroblasts and is identical
CC to the protein encoded by cyr61, an immediate early response gene
CC induced in quiescent BALB 3T3 cells by serum treatment. Residues
CC 49-56 of beta-IG-M1 conform to the GGCXXC motif reported in the
CC amino half of insulin-like growth factor (IGF) binding proteins.
CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
CC an amino acid sequence with strong homology to a motif found near the
CC C-terminal of the malarial circumsporozoite (CS) protein, which is
CC highly conserved among all species of malarial parasites sequenced

```

```

CC to date (designated region II). This motif is also found in
CC other proteins which have cell adhesive properties that mediate
CC cell-cell and cell-extracellular matrix interactions, such as
CC properdin, thrombospondin, and TRAP. The proteins encoded by
CC TGF-beta induced genes are likely to be involved in mediation of
CC the biological effects of TGF-beta relating to cell growth and
CC differentiation. See also R25566.
SQ Sequence 379 AA;

```

```

Query Match 41.0%; Score 1113; DB 5; Length 379;
Best Local Similarity 48.3%; Pred. No. 5.50e-95;
Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;

```

```

Db 1 mssstftrlava-vtllhl-trials-tcpaachcple-apkcapvgivrdgcgckvc 56
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 1 MTAASMGFVRVAFVLLALCSREAVGQNCSGPCRCDEPAPRCAGVSLVLDGCGCRVC 60
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 57 akqlnedcsktqpcdhtkglecnfgaestalkgicraqeegrpceynsriyngesafqn 111
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 61 AKQLGELCTERDPCDPKHGLFCDFGSPANRKGIVGTAK-DGAPCFGFTYRSGESFQS 119
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 117 ckhgctcidgavgcipicpqllelplnlgcpnprlvkvsqgceewcdedsikdlddd 176
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 120 SKKYQCTCLDGAAGVCMPLCSMDVRLPSPDCPPFRVRKLPKGCCEEWVCDK-PT 173
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 177 dlglidaseveltrnnelalgkgsalkrlpvgfptprvlfnplhahgkciavttswaq 236
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 174 -WVG--PA-----LA-----AY-R---LE--DTFGPDF-TM---IRAN--CLVQTWESA 208
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 237 cskcgtgistrvtndnpeclrvketricevrcpvgpvysslkkgkckskkspcvrf 296
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 209 CSKTCGMGISTRVNDNASCRLEKQSRLCMVRCCEADLEENIKKGGKIRTPKISKP 268
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 297 tyagcsavkkyrpkycgscvdrctplqtrtvkrmfrceedgemfaknmvmmiqscnyn 356
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 269 ELSGCTSMKTVRAKFCGVTGRCCTPHRTTTLTPVEFKPDGEVMMKNMFIKTCACHYN 328
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
Db 357 cphneasf-rly 368
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |
QY 329 CPGDNDI-FESLY 340
    |:::| | | | |:::| | | | |:::| | | | |:::| | | | |:::| | | | |

```

RESULT 6

```

ID R90919 standard; Protein; 375 AA.
AC R90919;
DT 25-JUN-1996 (first entry)
DE Connective tissue growth factor-2.
KW CTGF-2; connective tissue growth factor-2.
KW cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
KW insulin-like growth factor; fibroblast growth factor; Cry61.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= signal_peptide
FT Protein 25..375
FT /label= mature_protein
FN W09601896-A.
PD 25-JAN-1996.
PF 12-JUL-1994; U07736.
PR 12-JUL-1994; W0-U07736.

```


PA (HUMA-) HUMAN GENOME SCI INC.
PI Adams MD, Li H;
DR WPI; 96-097626/10.
DR N-PSDB; T12653.
PT Connective tissue growth factor-2 and DNA encoding it - useful to
PT enhance the repair of connective and support tissue, and to enhance
PT wound healing
PS Claim 1; Fig 1A-C; 46pp; English.
CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
CC (T12653) isolated from a human foetal lung cDNA library. The CTGF
CC polypeptides are structurally and functionally related to a family
CC of growth factors which include IGF (Insulin-like growth factor),
CC PDGF (platelet-derived growth factor), and FGF (fibroblast growth
CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
CC to Cry61. Cry61 is a growth factor-inducible immediate early gene
CC initially identified in serum-stimulated mouse fibroblasts. It encodes
CC a member of an emerging family of secreted proteins which are also a
CC group of cysteine-rich proteins. This group of GFs are important for
CC normal growth, differentiation, morphogenesis of the cartilaginous
CC skeleton of an embryo and cell growth.
SQ Sequence 375 AA;

Query Match 36.9%; Score 1002; DB 16; Length 375;
Best Local Similarity 47.2%; Pred. No. 3.88e-84;
Matches 175; Conservative 71; Mismatches 84; Indels 41; Gaps 21;

Db 6 vrelalvtllhl-trvqls-tcpadchcple-apkcapvgvlrvdgcgcckvcakqlne 62
Qy 9 VR-VAFVW-LTALCSRPVAVGQNSGCRPCDEPAPCPAGVSLVLDGCGCRCAKQLGE 66
Db 63 derktqpcdhtkglcnfgasatalkicraqsegrpcceynsriyqngsfqpnkqht 122
Qy 67 LCTERDPCDHPKGLGFCDFGSPANRKGIVCTAK-DGAPCIFGGTYRSGESFQSSCKYQCT 125
Db 123 cigwrrgacipclpaelapnlgcnpnrlvkvtdgcccceewvcdedsikdmedqdgllgk 182
Qy 126 CLDGAVG-CMPLCSMDVRLPSPDCFPFRRLKPGKCEEWVCD-----PK-DQT-VVGP 177
Db 183 glqfdaseveltrnnellavkggsalkzlpvfmgprilypnlggqkciqvgttswsqcsk 242
Qy 178 AL---AA---Y-R---L-E-----DF-----FGDPP-TM---IRA-NCLVQTTEWSACSK 211
Db 243 tegtgistrvtndnpeclrvketricevrpcgppvyslkgkckstkkskspvrfitya 302
Qy 212 TCGMGISTRVTDNASCRLKQSLCMVRPCEADLEENIKKGGKCIPTPKISKPIKPELS 271
Db 303 gclsvkkyrpkycgcvgdgrctpqltrtkvmrfpcdgetfsknvmmigskscnynpcph 362
Qy 272 GCTSMKTYRAKFCGVCTDGRCTPHRTTTLPEVFKPDGVEVMKMMFMFKTCAHYNCPG 331
Db 363 anea-afifyr 372
Qy 332 DNDIFESLYR 342

RESULT 7
ID R31608 standard; Protein; 205 AA.
AC R31608;
DT 24-MAY-1993 (first entry)
DE Homologous to chicken nov gene exon 3-4-encoded protein.

KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation; ss.
FN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
FI WPI; 93-036377/04.
DR Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 21; Page 39; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Fragment XXII (Q36044) is part of the 3rd and 4th exons of
CC the nov gene; nucleotide sequences which hybridise to Fragment XXII
CC under stringent conditions (i.e. 50% formamide, 5 x SSC) are claimed.
CC The claimed sequences preferably encode a protein with the sequence
CC XXIII (R31608) or an amino acid sequence 60% homologous to it.
SQ Sequence 205 AA;

Query Match 23.7%; Score 644; DB 6; Length 205;
Best Local Similarity 98.8%; Pred. No. 1.93e-49;
Matches 81; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 33 akdgapcifgtyvragcsfsgackvqctcldgavcmlpcmdvrlpsdpcpfrvkl 92
Qy 97 AKDGAPCIFGGTYRSGESFQSSCKYQCTCLDGA VCGMPLCSMDVRLPSPDCFPFRV 156
Db 93 pgkccceewvcddepkdgctvlgpa 114
Qy 157 PGKCCCEEWVCDPEKDTVVVGPA 178

RESULT 8
ID R31610 standard; Protein; 72 AA.
AC R31610;
DT 24-MAY-1993 (first entry)
DE Fragment XXVI with homology to chicken nov exon 3-encoded protein.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Homo sapiens.
FN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
FI WPI; 93-036377/04.
DR Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 25; Page 40; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived

CC probe. Nucleotide sequences of the invention include those which
CC code for a protein having at least 60% homology with the protein nov
CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
CC gene. In particular, the sequences of the invention encode a
CC protein having the sequence of Fragment XXVI (R31610). (Sequence
CC XXVI is described as a nucleotide sequence in the claims but is
CC correctly described as an amino acid sequence in the disclosure).
SQ Sequence 72 AA;

Query Match 21.3%; Score 579; DB 6; Length 72;
Best Local Similarity 97.2%; Pred. No. 3.25e-43;
Matches 70; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 dgapcifgtvyrsgesfqssckycctcldgavgcmlcsmdvrlpsdpcfprrvklpg 60
Qy 99 DGAPCIFGTVYRSGESFQSSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCFP RRVKLP G 158
Db 61 mceekwvcdpr 72
Qy 159 KCCEEWVCDPEK 170

RESULT 9

ID R31602 standard; Protein; 84 AA.
AC R31602;
DT 24-MAY-1993 (first entry)
DE Chicken nov protein fragment IX.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN WO9300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36035.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 8; Page 31, 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence was isolated from a gene bank prepared from
CC chicken embryonic fibroblasts screened with a tumour-derived probe.
CC Fragment VIII (Q36035) is derived from the 3rd. exon of the nov gene;
CC nucleotide sequences which hybridise to Fragment VIII under stringent
CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC sequences preferably encode a protein with at least 70% homology to
CC amino acid sequence IX (R31602) which is encoded by Fragment VIII.
SQ Sequence 84 AA;

Query Match 15.8%; Score 430; DB 6; Length 84;
Best Local Similarity 57.7%; Pred. No. 4.30e-29;
Matches 45; Conservative 19; Mismatches 13; Indels 1; Gaps 1;

Db 3 egdncvfdgmlyrngctfcpackycctcdgicgclprcnlglllpgdpcfporkievpg 62
Qy 99 DGAPCIFGTVYRSGESFQSSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCFP RRVKLP G 158

Db 63 eceekwvcd-prdevllg 79
Qy 159 KCCEEWVCDPEKQDTVVG 176

RESULT 10

ID R31609 standard; Protein; 71 AA.
AC R31609;
DT 24-MAY-1993 (first entry)
DE Encoded by chicken nov gene exon 3 fragment.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN WO9300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
DR WPI; 93-036377/04.
DR Nucleotide sequences hybridising to regions of chicken nov gene -
DR useful as probes for detecting complementary sequences to
DR evaluate development and/or differentiation of tumours
PS Claim 25; Page 40; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences of the invention include those which
CC code for a protein having at least 60% homology with the protein
CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
CC gene. (Sequence XXV is described as a nucleotide sequence in the
CC claims but is correctly described as an amino acid sequence in the
CC disclosure).
SQ Sequence 71 AA;

Query Match 15.2%; Score 412; DB 6; Length 71;
Best Local Similarity 59.7%; Pred. No. 2.09e-27;
Matches 43; Conservative 16; Mismatches 12; Indels 1; Gaps 1;

Db 1 egdncvfdgmlyrngctfcpackycctcdgicgclprcnlglllpgdpcfporkievpg 60
Qy 99 DGAPCIFGTVYRSGESFQSSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCFP RRVKLP G 158
Db 61 eceekwvcd-pr 71
Qy 159 KCCEEWVCDPEK 170

RESULT 11

ID R31603 standard; Protein; 84 AA.
AC R31603;
DT 24-MAY-1993 (first entry)
DE Polypeptide X homologous to chicken nov protein fragment.
DE avian nephroblastoma; avian myeloblastoma virus;
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Homo sapiens.
PN WO9300430-A.
PD 07-JAN-1993.

25-JUN-1992; F00589.
25-JUN-1991; FR-007807.
(CNRS) CENT NAT RECH SCI.
Martinerie C, Perbal B;
WPI: 93-036377/04.
N-FSDB; Q36036.
Nucleotide sequences hybridising to regions of chicken nov gene -
useful as probes for detecting complementary sequences to
evaluate development and/or differentiation of tumours
Claim 10; Page 31; 67pp; French.
Nucleotide sequences which hybridise to Fragment VIII (Q36035) of
the 3rd. exon of chicken nov gene under stringent conditions (i.e.
50% formamide, 5 x SCC) are claimed. The claimed sequences preferably
encode a protein with at least 70% homology to R31602 which is encoded
by Fragment VIII. Such nucleotide sequences are further
characterised in that they comprise at least part of a 700bp pSti
fragment derived from a recombinant clone isolated from a human
placenta DNA library and encode the amino acid sequence of
Fragment X (R31603).
Sequence 84 AA;
SO

Query Match	14.9%	Score 405;	DB 6;	Length 84;
Best Local Similarity	53.7%	Prod. No. 9.43e-27;		
Matches	44;	Conservative	17;	Mismatches 21;
			Indels	0;
			Caps	0;
Db	3	egdnvcfdgviyrsekfqpsekfqtctrdgqigevprcqldvlvpncpaprkvvevp	62	
	:	: : : : : : : : : : : : :		
Qy	99	DGAPCIFGTVYRSGESFQSSCRYQCTCLDGAVGCMPLCSMDVRLPSPDCCFFPRRVKLP	158	
Db	63	eccekwicgpgdeedslggtla	84	
	:	: : : : : : : : : : : :		
Qy	159	KCEEWVCDEPKDQTVVGPALA	180	

RESULT	12	
ID	R46078	standard; Protein; 124 AA.
AC	R46078;	
DT	19-OCT-1994	(first entry)
DE	CYR61 like protein.	
DE	Human cDNA; library; enzyme; protein.	
KE	Homo sapiens.	
OS	WO9403599-A.	
PN	PN	
PD	17-FEB-1994.	
PF	04-AUG-1993;	J01095.
PF	04-AUG-1992;	JP-208077.
PP	13-NOV-1992;	JP-327619.
PR	26-FEB-1993;	JP-061431.
PPA	(SAGA) SAGAMI CHEM RES CENTRE.	
PPF1	Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;	
PPF2	WPI; 94-065688/08.	
DR	N-PSDB; Q57417.	
PPPT	cDNA of human origin and proteins coded by it - which may be	
PPPT	expressed by in vivo or in vitro translation using sense RNA or	
PPPT	antisense DNA corresponding to the cDNA.	
PPS	Claim 1; Page 32-33; 167pp; Japanese.	
CCCC	mRNA expressed in human fibroarcoma cell line HT-1080 was	
CCCC	isolated and used to construct a cDNA library using vector	
CCCC	pKAL. Clone HP00021 encoding CYR61-like protein	
CC	was isolated.	
CC	Sequence	124 AA;
ISQ		

	Query Match	14.6%	Score 395;	DB 10;	Length 124;
	Best Local Similarity	58.4%	Pred. No. 8.10e-26;		
	Matches	45;	Mismatches 16;	Indels 1;	Gaps 1;
Db	6	gicraqeagrpceynsriyngesfgpnckhctctidgavgcipclpqelslplngcpnp	65		
		: : : : : : : : : : : : : : : :			
Qy	93	GVCTAK-DGAPCIFGTVYRSGESFQSCYQCTCIDGAVGCMPLCSMDVRLPSDPCFP	151		
Db	66	rlvkvtggccceewvcde	82		
		: : : : : : : : : : : : : : : :			
Qy	152	RRVKLPKGKCEEWVCDE	168		

RESULT 13
ID R31605 standard; Protein: 72 AA.

AC R31603;
 DT 24-MAY-1993 (first entry)
 DE Homologous to chicken nov exon 4-encoded protein.
 DE avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation; CTGF.

07-JAN-1993. PD
25-JUN-1992; PF F00589. PD
25-JUN-1991; FR-007807. PR
(CNRS) CENT NAT RECH SCI. PA
Martinerie C, Perbal B; PI
WPI; 93-036377/04. PI
N-FSDB; Q36039. DR
Nucleotide sequences hybridising to regions of chicken nov gene -
useful as probes for detecting complementary sequences to PPT
evaluate development and/or differentiation of tumours PPT
Claim 14; Page 34; 67pp; French. PPT
The chicken nov gene is stimulated in avian nephroblastoma induced
by avian myeloblastoma virus but not in normal adult kidney. A CCC
1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared CCC
from chicken embryonic fibroblasts screened with a tumour-derived CCC
probe. Fragment XIII (Q36038) is derived from the 4th. exon of the CCC
nov gene; nucleotide sequences which hybridise to fragment XIII CCC
under stringent conditions (i.e. 50% formamide, 5 x SCC) are claim CCC
The claimed sequences preferably encode a protein with at least 86 CCC
homology to R31604 which is encoded by fragment XIII. Pref. the CCC
sequence of the invention comprises a sequence which encodes amino CCC
acid sequence XV (R31605). CCC
Sequence 72 AA; SQ

			Query Match	11.6%; Score 314; DB 6; Length 72;
			Best Local Similarity	58.3%; Pred. No. 2.49e-18;
			Matches:	Conservative 11; Mismatches 19; Indels 0; Gaps 0;
Db	l	ayrpeatlgvevdsavncieqtetwacsksccgmfgfstrvtvrnrcgcmklkqrlcmvr	60	
	:	:	: : : : : :	: : :
Oy	181	AYRLDTFGDPTMIRANCLVQTWEASCKTCGNGI STRVTNDNASCRLEQSRLCMYR	240	
Db	61	pceqepeeqptdk	72	
Oy	241	PCEADLENIKK	252	

RESULT 14

ID R31604 standard; Protein; 70 AA.
AC R31604;
DT 24-MAY-1993 (first entry)
DE Chicken nov protein fragment XIV.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36038.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 13; Page 34; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence was isolated from a gene bank prepared from
CC chicken embryonic fibroblasts screened with a tumour-derived probe.
CC Fragment XIII (Q36038) is derived from the 4th. exon of the nov gene;
CC nucleotide sequences which hybridise to Fragment XIII under stringent
CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC sequences preferably encode a protein with at least 86% homology to
CC amino acid sequence XIV (R31604) which is encoded by Fragment XIII.
SQ Sequence 70 AA;

Query Match 11.5%; Score 312; DB 6; Length 70;
Best Local Similarity 64.6%; Pred. No. 3.79e-18;
Matches 42; Conservative 9; Mismatches 14; Indels 0; Gaps 0;
Db 1 avrqeatlqdvdsasanciegttewasacskcgmgtfstrvtrnngqcmvkvqtricmmr 60
QY 181 AYRLEDTFGPDFTMIRANCLVQTTEWSACSKTCGIGISTRVTNDNASCRLEKOSRLCMVR 240
Db 61 pcene 65
QY 241 PCEAD 245

RESULT 15
ID R31601 standard; Protein; 75 AA.
AC R31601;
DT 24-MAY-1993 (first entry)
DE Chicken nov protein fragment V.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to

PT evaluate development and/or differentiation of tumours
PS Claim 5; Page 28; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences which hybridise to Q36031 or specified
CC sub-fragments of it, under stringent conditions (i.e. 50% formamide,
CC 5 x SCC), are claimed. The claimed sequences preferably encode a
CC protein with amino acid sequence V (R31601).
SQ Sequence 75 AA;

Query Match 10.2%; Score 278; DB 6; Length 75;
Best Local Similarity 48.0%; Pred. No. 4.66e-15;
Matches 36; Conservative 14; Mismatches 22; Indels 3; Gaps 2;
Db 2 aatgcpqpqgpcratp-ptcapgvrvldgscclvcargqscsclpcdcassgly 60
QY 24 AVGNCSGPC--RCPDEPAPRCPAGVSLVLDGGCCRCVCAKQIGELCTERDFCDPHKGLF 81
Db 61 cdreadpsnqtgict 75
QY 82 CDFGSPANRKIGVCT 96

Search completed: Wed Sep 17 09:36:20 1997
Job time : 72 secs.

maryh@stic

4380-1

NeWSprinter20

Wed Sep 17 11:17:00 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

 T V A P S R E L H
 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by IntelliGenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Wed Sep 17 09:33:28 1997; MacPar time 13.07 Seconds
 Tabular output not generated.
 771.467 Million cell updates/sec

Title: >05-08-167-628-2
 Description: (1-349) from 5408040.ppe
 Perfect Score: 2713
 Sequence: 1 MTAASGMPVRVAFVILLALC.....PGNDIFESLYRMYGDMA 349

Scoring table:
 PAM 150
 Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pir51
 1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
 8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
 13:unann9 14:unann10 15:unann16:unrev

Statistics: Mean 42.720; Variance 82.202; scale 0.520

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	2713	100.0	349	13	connective tissue gr	0.00e+00
2	2521	92.9	348	14	beta IG-M2 protein p	0.00e+00
3	2514	92.7	348	14	fisp-12 protein prec	0.00e+00
4	1460	53.8	351	13	NOV protein - chicke	3.42e-284
5	1415	52.2	357	13	gene novH protein -	2.53e-274
6	1152	42.5	375	13	CEF-10 protein precu	7.43e-217
7	1113	41.0	379	14	gene Cfr61 protein p	2.27e-208
8	133	4.9	254	14	insulin-like growth	1.32e-06

9	133	4.9	254	14	I48599	insulin-like growth	1.32e-06
10	128	4.7	271	13	JC4584	insulin-like growth	8.02e-06
11	125	4.6	271	14	JC1463	insulin-like growth	2.33e-05
12	126	4.6	835	13	JP0076	nel protein - chicke	1.63e-05
13	123	4.5	258	13	B37252	insulin-like growth	4.71e-05
14	121	4.5	455	6	A43918	TGF-beta-related pro	9.49e-05
15	116	4.3	254	14	I48603	insulin-like growth	5.32e-04
16	117	4.3	277	13	I52825	gene MAC25 protein -	3.78e-04
17	117	4.3	282	13	S50031	prostaglandin-stimula	3.78e-04
18	118	4.3	1348	12	A43917	probable epidermal c	2.68e-04
19	118	4.3	1348	12	S27812	probable epidermal c	2.68e-04
20	115	4.2	272	6	A36082	insulin-like growth	7.48e-04
21	113	4.2	2813	3	VWHD	von Willebrand facto	1.47e-03
22	110	4.1	263	6	S23009	insulin-like growth	4.01e-03
23	111	4.1	271	14	I48604	insulin-like growth	2.88e-03
24	111	4.1	304	6	A33274	insulin-like growth	2.88e-03
25	111	4.1	305	14	I48601	insulin-like growth	2.88e-03
26	111	4.1	305	6	JN0508	insulin-like growth	2.88e-03
27	110	4.1	3020	13	A43932	mucin 2 precursor, i	4.01e-04
28	109	4.0	124	6	S25113	insulin-like growth	5.39e-03
29	109	4.0	272	14	I48600	insulin-like growth	5.39e-03
30	106	3.9	258	14	A45403	insulin-like growth	1.49e-02
31	105	3.9	272	13	A53748	insulin-like growth	2.07e-02
32	106	3.9	807	14	A38152	F-spondin - rat	1.49e-02
33	102	3.8	259	2	IOH01	insulin-like growth	5.40e-02
34	102	3.8	291	2	JN0064	insulin-like growth	5.40e-02
35	104	3.8	310	6	A60967	insulin-like growth	2.85e-02
36	103	3.8	328	2	A41927	insulin-like growth	3.93e-02
37	103	3.8	500	4	P1WL6	L1 protein - human p	3.93e-02
38	102	3.8	798	7	A40526	integrin beta-7 chai	5.40e-02
39	104	3.8	3133	16	S70920	lectin - silkworm	2.85e-02
40	104	3.8	3133	16	S52093	circumsporozoite pro	7.42e-02
41	101	3.7	388	7	A39756	prepro osteogenic pro	7.42e-02
42	101	3.7	402	6	A45056	circumsporozoite pro	7.42e-02
43	101	3.7	405	7	S05428	circumsporozoite pro	7.42e-02
44	101	3.7	412	3	OZ20AF	circumsporozoite pro	7.42e-02
45	101	3.7	424	12	A54533	circumsporozoite pro	7.42e-02

ALIGNMENTS

RESULT	1
ENTRY	A40551
TITLE	connective tissue growth factor - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 25-Aug-1995
ACCESSIONS	A40551; S44205
REFERENCE	A40551
#authors	Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
#journal	J. Cell Biol. (1991) 114:1285-1294
#title	Connective tissue growth factor: a cytosine-rich mitogen secreted by human vascular endothelial cells is related to the SRC-induced immediate early gene product CEF-10.
#cross-references	WUID:91373462
#accession	A40551
#molecule_type	mRNA
#residues	1-349
#label	BRA
#authors	S44205 Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.;

submitters, T.F.
submitted to the EMBL Data Library, April 1994
Differential cloning and expression of human connective
tissue growth factor.

accession S44205 preliminary
#status preliminary
#molecule_type mRNA
#residues 1-349 ##label OEM
#cross-references EMBL:X78947
#length 349 #molecular-weight 38069 #checksum 8930

Query Match 100.0%; Score 2713; DB 13; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mtaasmgprvrvafvllalcalcarpavqncsgpcrcpdepaprcpavagvslvldgcgcrcvc 60
Qy 1 MTAASMGPRVRAVFVLLALCSRPVAVGQNCSGPCRCDEPAPRCFAGVSLVLDGCGCCRCVC 60

Db 61 akqlgelcterdpdphkgldfcdgspgpanrkigvctakdgapciifgtvyrsgesfqsac 120
Qy 61 AKQLGELCTERDPCDPHKGLFCDGSPGPNRKIGVCTAKDGPAPCIIFGTVYRSGESFQSSC 120

Db 121 kyqctcldgavgcplcamdvrlpdpdpfprvrklpgkceewvcdpdkdtvvgpala 180
Qy 121 KYQCTCLDGA VGCPLCAMDVRLPDPDPFPRVRKLPKCEEWVCDPDKDTVVGPA LA 180

Db 181 avrledtfgpdpmtirancivqtetwascsktcgmgiatrvtdnnaascrlkgsrlcmvr 240
Qy 181 AVRLEDTFGDPMTIRANCLVQTETWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVR 240

Db 241 pceadleenikkkgkcirtkpkiskpikfelsgctsmktyrakfcgvctdgrcctphrttt 300
Qy 241 PCEADLEENIKKGKCIKTPKISKPIKELSGCTSMKTYRAKFCGVCTDGRCTPHRTTT 300

Db 301 lpvefkcpdgvmmkmmfiktacachyncpogndndifeslyyrkmygdma 349
Qy 301 LPVEFKCPDGEVMKMMFIKTCACHYNCPOGNDNDFESLYYRMYGDMA 349

RESULT 2
ENTRY beta IG-M2 protein precursor - mouse
TITLE #formal_name Mus musculus #common name house mouse
ORGANISM 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change
DATE 30-Sep-1993

ACCESSIONS A40578
REFERENCE A40578
#authors Brunner, A.; Chinn, J.; Neuberger, M.; Purchio, A.F.
#journal DNA Cell Biol. (1991) 10:293-300
#title Identification of a gene family regulated by transforming
growth factor-beta.
#cross-references MIM:91229699

accession A40578 preliminary
#status preliminary
#molecule_type mRNA
#residues 1-348 ##label BRU
#length 348 #molecular-weight 37794 #checksum 9188

Query Match 92.9%; Score 2521; DB 14; Length 348;

Best Local Similarity 91.1%; Pred. No. 0.00e+00;
Matches 318; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

Db 1 mlasavagpiialv-llalettrpatgdcacqcaaeapahcpagvalvidgcgcrcvc 59
Qy 1 MTAASMGPRVRAVFVLLALCSRPVAVGQNCSGPCRCDEPAPRCFAGVSLVLDGCGCCRCVC 60

Db 60 akqlgelcterdpdphkgldfcdgspgpanrkigvctakdgapcvfggavragesfqsac 119
Qy 61 AKQLGELCTERDPCDPHKGLFCDGSPGPNRKIGVCTAKDGPAPCIIFGTVYRSGESFQSSC 120

Db 120 kyqctcldgavgcplcamdvrlpdpdpfprvrklpgkceewvcdpdkdtvvgpala 179
Qy 121 KYQCTCLDGA VGCPLCAMDVRLPDPDPFPRVRKLPKCEEWVCDPDKDTVVGPA LA 180

Db 180 avrledtfgpdpmtirancivqtetwascsktcgmgiatrvtdnndifeslyyrkmygdma 239
Qy 181 AVRLEDTFGDPMTIRANCLVQTETWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVR 240

Db 240 pceadleenikkkgkcirtkpkiskpikfelsgctsmktyrakfcgvctdgrcctphrttt 299
Qy 241 PCEADLEENIKKGKCIKTPKISKPIKELSGCTSMKTYRAKFCGVCTDGRCTPHRTTT 300

Db 300 lpvefkcpdgvmmkmmfiktacachyncpogndndifeslyyrkmygdma 348
Qy 301 LPVEFKCPDGEVMKMMFIKTCACHYNCPOGNDNDFESLYYRMYGDMA 349

RESULT 3
ENTRY fisp-12 protein precursor - mouse
TITLE #formal_name Mus musculus #common name house mouse
ORGANISM 19-May-1994 #sequence_revision 19-May-1994 #text_change
DATE 19-May-1994

ACCESSIONS A53228
REFERENCE A53228
#authors Rybeck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
#journal Cell Growth Differ. (1991) 2:225-233
#title Structure, mapping, and expression of fisp-12, a growth
factor-inducible gene encoding a secreted cysteine-rich
protein.

accession A53228 preliminary
#status preliminary
#molecule_type DNA
#residues 1-348 ##label RYS
#cross-references GB:M70641

GENETICS fisp-12
#gene fisp-12
#length 348 #molecular-weight 37793 #checksum 9470

Query Match 92.7%; Score 2514; DB 14; Length 348;
Best Local Similarity 90.8%; Pred. No. 0.00e+00;
Matches 317; Conservative 21; Mismatches 10; Indels 1; Gaps 1;

Db 1 mlasavagpiialv-llalettrpatgdcacqcaaeapahcpagvalvidgcgcrcvc 59
Qy 1 MTAASMGPRVRAVFVLLALCSRPVAVGQNCSGPCRCDEPAPRCFAGVSLVLDGCGCCRCVC 60

Db 60 akqlgelcterdpdphkgldfcdgspgpanrkigvctakdgapcvfggavragesfqsac 119
Qy 61 AKQLGELCTERDPCDPHKGLFCDGSPGPNRKIGVCTAKDGPAPCIIFGTVYRSGESFQSSC 120

```
QY 61 AKQLGELCTERDPCDPHKGLFCDFGSPANRKGIVCTAKDGAPCIFGCTVYRSGESFQSSC 120
Db 120 kvqctcldgavgcvpclcamdvrlpspcdpfprvrklpqkckewcvddepkdrtaavgpala 179
QY 121 KYQCTCLDGAAGVCMPLCSMDVRLPSPCDPFPRVRKLPKGCCEEWVCDPEKDPQTWGPALA 180
Db 180 avrledtfgpdpmmranclvqtteasacsktcmgmistrvntndntfcrlckqaricmvr 239
QY 181 AYRLEDTFGPDPMTIRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVR 240
Db 240 pceadleenikgkciirtpkiakpvkfelagctavktyrakcgvctdgrctphrttt 299
QY 241 PCEADLEENIKGKCIIRTPKISPKIFELSGCTSMKTYRAKFCGCTDGRCTPHRTTT 300
Db 300 lpfvfkcpdgelmknmfiktacachncpogndndifsealyrkmvgdma 348
QY 301 LPVEFKCPDGEVMKKNMFIKTCACHYNCPGNDIFESLYRKMVGDMA 349

RESULT 4
ENTRY #20078 #type complete
TITLE NOV protein - chicken
ORGANISM #formal name Gallus gallus #common name chicken
DATE 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change
10-Nov-1995

ACCESSIONS S20078
REFERENCE #authors Jolliot, V.; Martinerie, C.; Dambrie, G.; Plassiart, G.;
#journal Brisac, M.; Crochet, J.; Perbal, B.
#title Mol. Cell. Biol. (1992) 12:10-21
#cross-references EMBL:X59284
#accession S20078
#status preliminary
#molecule type mRNA
#residues 1-351 #label JOL
#cross-references EMBL:X59284

GENETICS
#gene NOV
SUMMARY #length 351 #molecular-weight 38268 #checksum 4069

Query Match 53.8%; Score 1460; DB 13; Length 351;
Best Local Similarity 54.2%; Pred. No. 3.42e-284;
Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;

Db 3 tgggggplvllllllllrpceavsgreaacprcpggrcpaep-prcagpvpavldgcgccl 61
QY 2 TAASMG-SVRVAVVLLALCSRPAGVQNCSPC--RCPEPAPRCAGVSLVLDGCGCCR 58
Db 62 vcarqgscapllpcdesaglycdrpdegagagcmvlegdncvfdgmlyrngctfcp 121
QY 59 VCAKQLGELCTERDPCDPHKGLFCDFGSPANRKGIVCTAKDGAPCIFGCTVYRSGESFQSS 118
Db 122 ackvqctcdgqigclpncnlgllpgdpdpfprkicvpgcecekwcd-prdevllgff 180
QY 119 SKYQCTCLDGAAGVCMPLCSMDVRLPSPCDPFPRVRKLPKGCCEEWVCDPEKDPQTWGP- 177
Db 181 amaaayrqeatlgidvdsasancieqtteasacskcmgmfstvtrnrrnqgcmvktqrlc 240
```

```
QY 178 ALAAYLEDTFGDPDMIRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLC 237
Db 241 mmrppcene-epsdkkgkciqtckkamkavrfeykncvatvtykprycgclndgrrctphn 299
QY 238 MVRPEADLEENIKGKCIIRTPKISPKIFELSGCTSMKTYRAKFCGCTDGRCTPHR 297
Db 300 tktiqvfcropgqgkflkppmmlintevchgnpcqannaffqpl 342
QY 298 TTTLPVEFKCPDGEVMKKNMFIKTCACHYNCPGNDIFESLYRKMVGDMA 339

RESULT 5
ENTRY #138069 #type complete
TITLE gene novH protein - human
ORGANISM #formal name Homo sapiens #common name man
DATE 17-May-1996 #sequence_revision 17-May-1996 #text_change
17-May-1996

ACCESSIONS I38069
REFERENCE #authors Martinerie, C.; Huff, V.; Joubert, I.; Badzioch, M.;
#journal Saunders, G.; Strong, I.; Perbal, B.
#title Oncogene (1994) 9:2729-2732
#cross-references EMBL:X78351; NID:g587422; CDS_PID:g825696
#accession I38069
#status preliminary; translated from GB/EMBL/DBJ
#molecule type DNA
#residues 1-357 #label RES
#cross-references EMBL:X78351; NID:g587422; CDS_PID:g825696

GENETICS
#introns 28/3; 104/1; 188/1; 259/3
#note gene name novH
SUMMARY #length 357 #molecular-weight 39162 #checksum 1640

Query Match 52.2%; Score 1415; DB 13; Length 357;
Best Local Similarity 51.8%; Pred. No. 2.53e-274;
Matches 177; Conservative 68; Mismatches 91; Indels 6; Gaps 5;

Db 18 ltfll-llhllgvaatqrcpqpgrcpatp-ptcagpvrvavldgscclvcargesc 75
QY 11 VAFVLLALCSRPAGVQNCSPC--RCPEPAPRCAGVSLVLDGCGCCRCAKQLGELC 68
Db 76 sdlepdeasaglycdraadpsnqgtgictavegdncvfdgviyrsgkfqackfctord 135
QY 69 TERDPCDPHKGLFCDFGSPANRKGIVCTAKDGAPCIFGCTVYRSGESFQSSCKYQCTCLD 128
Db 136 gqigcvproqlavllpepncpaprkvevpgcecekwicgpxdeedalgltlaayrpeatl 195
QY 129 GAVGCMPLCSMDVRLPSPCDPFPRVRKLPKGCCEEWVCDPEKDPQTWGPALAAVRLDTF 188
Db 196 gvevasvsnclieqtteasacskcmgmfstvtrnrrnqcmktrclcmvrpceqeq 255
QY 189 GPDPMTIRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVRPEADLEE 248
Db 256 ptdkkgkclrrkkalkaihlgfknctalhtykrpfcgvcadgrrctphntktiqaeffc 315
QY 249 NI-KKGGKCIIRTPKISPKIFELSGCTSMKTYRAKFCGCTDGRCTPHRTTLPVEFKC 307
```



```

Db   316 spqgiavkpvvmvigtcthtnnpknneaflelelkttrgkm 357
      : ::::| | | | | | | | | | | | | | | | | | | | | |
Qy   308 PDGEVMKKNNMFIKTCACHYNCPGNDIP-ESLYRRMYGDM 348
      : ::::| | | | | | | | | | | | | | | | | | | | | |

RESULT          6
ENTRY           A14128             #type complete
TITLE          CF9-10 protein precursor - chicken
ORGANISM       Gallus gallus gallus #common name chicken
DATE           03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change
              05-May-1993
ACCESSIONS     A14128
REFERENCE      A14128
#authors       Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
#journal       Proc. Natl. Acad. Sci. U.S.A. (1989) 86:1178-1182
#title         Identification of a phorbol ester-repressible v-src-inducible
               gene.
#cross-references MDID:89145206
#accession     A14128
##status       preliminary
##molecule type mRNA
##residues     1-375 ##label SIM
##cross-references GB:J04496
#length        375 #molecular-weight 40651 #checksum 1417

SUMMARY

Query Match      42.5%; Score 1152; DB 13; Length 375;
Best Local Similarity 48.5%; Pred.No. 7.43e-217;
Matches 176; Conservative 76; Mismatches 77; Indels 34; Gaps 19;

Db   7 ipalaal-lclarlalgpccapvcqba-aapcagpgvlvpdgccckvcakqlined 64
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy   10 RVAFVLLALC-SRPAGVGNCSGPCRCDEPAPPCPVGVSLVDGCCRCRCAKQLGELC 68
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db   65 strtqcptdkgtiecnfgaspaatngicraqegrpceynskiyngesfqpncqhgtci 124
      : ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Qy   69 TERDPCDPHKGLFCDFGSPANKIGVCTAK-DGAPCIFGGTVRSGESFSQSSCKYQCTCL 127
      : ||||| | | | | | | | | | | | | | | | | | | | | | | | |

Db   125 dgavgcipicpqelslpnlgcpsprlvkvpgqccsewcvdeskdaleelgeffskefgid 184
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy   128 DGAVGCMPLCSMDVRLPSPDCFPFRARVLKPKCCEEWCDEPKD--QTVVVG--P---AL- 179
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db   185 asegetrnnellaivkgklmpvfgspegsrafenpkciavqttswsqscsktcgtgist 244
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy   180 AA---Y-R--I-E-----TFGDPP-TM-I-RANCLVQTTEWSACKTCGMGIST 219
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db   245 rvtndnpdeklketricevrpcgqpsyaslkqkktktkkspsvrftyagcsavkky 304
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy   220 RVTNDNASCRLEKQSRLCMVRPEADLEENIKKGKKCIPTPKISKPIKFELSGETSMKVY 279
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db   305 rpkygcsgvdgrcctqrtrtkirfrdddetfkusvmmiqsercny nophaney-p- 362
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy   280 RAKFCGCVTDGRCCTPHRTTTLPLVEFKCPDGVMKKNMFFIKTCACHYNCPGNDIFESL 339
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db   363 fyr 365
      : | |
Qy   340 YIR 342

RESULT          7
ENTRY           A35669             #type complete

```

gene	CYR61 protein precursor - mouse
formal name	Mus musculus #common name house mouse
28-Sep-1990	#sequence_revision 18-Nov-1992 #text_change
18-Oct-1996	
A35669	I48319
A35669	
O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.	
Mol. Cell. Biol. (1990) 10:3569-3577	
Expression of cyr61, a growth factor-inducible	
immediate-early gene.	
#cross-references MUID:90287146	
accession	A35669
##status	preliminary
##molecule_type	mRNA
##residues	1-379 #label OAB
##cross-references	GB:M32490
##note	the authors translated the codon GAT for residue 337
	Gln
I48319	
Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.	
Nucleic Acids Res. (1991) 19:3261-3267	
Promoter function and structure of the growth	
factor-inducible immediate early gene cyr61.	
#cross-references MUID:91288203	
accession	I48319
##status	translated from GB/EMBL/DBJ
##molecule_type	DNA
##residues	1-379 #label RES
##cross-references	EMBL:X56790; NID:g50632; CDS_PID:g50633
GENETICS	
CYR61	
#gene	
#introns	217/3; 9371; 208/1; 279/3
SUMMARY	
length	379 #molecular-weight 41709 #checksum 3726
Query Match	41.0%; Score 1113; DB 14; Length 379;
Best Local Similarity	48.3%; Pred. No. 2,27e-208;
Matches	180; Conservative 79; Mismatches 76; Indels 38; Gaps 2
Db	1 masstfrtlava-vtlhl-trials-topaachople-apkcapgvglvrdgcgcckvc 56
QY	1 MTAASMGFVRVAVLLALCSRPVQNGSGFCRCPDEPAPCPAGVSILVDGGCGRVC 60
Db	57 aqlnedcsktqpcdhtkglecnfgasalkgicraesgrpceynsrlyqngesfqn 116
QY	61 AKQLGELCTEDPCDPHGKFLCDFGSPANRKTGVGTAK-DGAPCIFGTVYRSGFSQS 119
Db	117 ckhqtctidgavgpiclpqelqsalpnlgcmprlvkvsgqceewcedsidslddgd 176
QY	120 CKYQCTCLDGVAGVGMFLCSMDVLESPDCFPFRVRVKLPGKCEEWCODE-P-KD--QT 173
Db	177 dlglidaeveeltrnneliaigksgsalrklrvfgtgeprvlnflhahgqtkcivqttswaq 236
QY	174 -VVG--PA---LA-----AY-R---LE--DTFGDP--TM---IRAN---CLVQTTEWSA 208
Db	237 cskcsgtgiatrvtndnpeclrvketricevrcpgpyysalktkkctktskpeprvf 296
QY	209 CSKTCGMGISITRVTDNASCRLEKSRQGMVRCPEADLEENIKKGGKCIHTPKISKPIK 268
Db	297 tyagcskvkrypkvgqscvdgrctptlqtrrvkmrfcedgmfsknvmmigscsknyn 356

Qy 269 ELSGCTSMKTYRAKFCVCTDGRCTPHRTTTLPEFKCPDGEVMKNNMFIKTCACHYN 328

Db 357 cphpneaf-rlly 368
|| | : | | |

Qy 329 CPGNDI-FESLY 340

RESULT 8
ENTRY #type complete
TITLE insulin-like growth factor-binding protein 4 precursor - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995

ACCESSIONS JCI1464; A37252; E40403; A61120; B33570; A61118
REFERENCE JCI1464
#authors Gao, L.; Ling, N.; Shimasaki, S.
#journal Biochem. Biophys. Res. Commun. (1993) 190:1053-1059
#title Structure of the insulin-like growth factor binding protein-4 gene.

#accession JCI1464
#molecule_type DNA
#residues 1-254 #label GAO
#cross-references GB:L08276

REFERENCE A37252
#authors Shimasaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.
#journal Mol. Endocrinol. (1990) 4:1451-1458
#title Molecular cloning of the cDNAs encoding a novel insulin-like growth factor-binding protein from rat and human.

#cross-references MUID:91133415
#accession A37252
#molecule_type mRNA
#residues 1-254 #label SHI

REFERENCE A40403
#authors Shimasaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
#journal J. Biol. Chem. (1991) 266:10646-10653
#title Identification of five different insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human.

#cross-references MUID:91244847
#accession E40403
#molecule_type protein
#residues 22-26,'X',28-29,'X',31-37,'X',39-40 #label SH2

REFERENCE A61120
#authors Ceda, G.P.; Fielder, P.J.; Henzel, W.J.; Louie, A.; Donovan, S.M.; Hofman, A.R.; Rosenfeld, R.G.
#journal Endocrinology (1991) 128:2815-2824
#title Differential effects of insulin-like growth factor (IGF)-I and IGF-II on the expression of IGF binding proteins (IGFBPs) in a rat neuroblastoma cell line: isolation and characterization of two forms of IGFBP-4.

#accession A61120
#molecule_type protein
#residues 'X',23-26,'X',28-29,'X',31-37,'X',39-43,'X',45,'X',47-48 #label CED
#note both glycosylated and nonglycosylated forms of this protein were observed

REFERENCE A33570
#authors Shimonaka, M.; Schroeder, R.; Shimasaki, S.; Ling, N.
#journal Biochem. Biophys. Res. Commun. (1989) 165:189-195
#title Identification of a novel binding protein for insulin-like

#cross-references MUID:90073708
#accession B33570
#molecule_type protein
#residues 22-26,'X',28-29,'X',31-37,'X',39-43,'X',45-52,'X',54,'XX',57-58,'X',60-61 #label SH3

REFERENCE A61118
#authors Cheung, P.T.; Smith, E.P.; Shimasaki, S.; Ling, N.; Chernausek, S.D.
#journal Endocrinology (1991) 129:1006-1015
#title Characterization of an insulin-like growth factor binding protein (IGFBP-4) produced by the B104 rat neuronal cell line: chemical and biological properties and differential synthesis by sublines.

#accession A61118
#molecule_type protein
#residues 22-26,'X',28-29,'XX',32-33 #label CHE

GENETICS 117/1; 165/3; 210/3
#superfamily thyroglobulin type I repeat homology glycoprotein

CLASSIFICATION 1-21
#domain signal sequence #status experimental #label SIG

FEATURES 22-254
#product insulin-like growth factor-binding protein 4
170-245
#status experimental #label MAY
#domain thyroglobulin type I repeat homology #label THY1\

125
#binding site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 254 #molecular-weight 27745 #checksum 4589

Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No. 1.32e-06;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;

Db 1 mlpfglvaalllaagrpelgde-aihcpcseeklarcpvgeelvrpggceate 59
| | : | | | : | | | : | | : | | : | | : | | | | | | | | | |
Qy 6 MGPVRVAFVLLALCSRPVAGQNCSGPCR-CPDEPAPRC-PA-GVS-IVLD-GGCCRVC 60
| | : | | | : | | | : | | : | | : | | : | | | | | | | | | |
Db 60 alglgmpe 67
| | : | | | : | | | : | | : | | : | | : | | | | | | | | | |
Qy 61 AKQLGELC 68

RESULT 9
ENTRY #type complete
TITLE insulin-like growth factor binding protein 4 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996

ACCESSIONS I48599; S38668
REFERENCE I48599
#authors Bethel, C.R.; Vitullo, J.C.; Miller, R.E.; Aron, D.C.
#journal Biochem. Mol. Biol. Int. (1994) 34:385-392
#title Molecular cloning of mouse insulin-like growth factor binding protein 4 (IGFBP4) cDNA and expression of a fusion protein with IGF-binding activity.

#cross-references MUID:95152444
#accession I48599
#status preliminary; translated from GB/EMBL/DBJ

```

##molecule_type DNA
##residues 1-254 ##label RES
##cross-references EMBL:X76066; NID:g416032; CDS PID:g416033
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE
170-245 #domain thyroglobulin type I repeat homology #label THY1
SUMMARY
length 254 #molecular-weight 27807 #checksum 3509
Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No. 1.32e-06;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
Db 1 mlpfglvaallaaqprslgde-aihhpcseeklarcrppvgceelvrepqgcgcac 59
Qy 6 MGFVRVAFVLLALCSRPVAVGQNSGFC-CPDEPAPRC-PA-GVSLVLD-GCGCCRCVC 60
Db 60 alqlmpc 67
Qy 61 AKQLGELC 68

RESULT 10
ENTRY JC4584 #type complete
TITLE insulin-like growth factor binding protein-5 - pig
ORGANISM #formal name Sus scrofa domestica #common name domestic pig
DATE 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 03-Jun-1996
ACCESSIONS JC4584
REFERENCE JC4584
#authors White, M.E.; Diao, R.; Hathaway, M.R.; Mickelson, J.; Dayton, W.R.
#journal Biochem. Biophys. Res. Commun. (1996) 218:248-253
#title Molecular cloning and sequence analysis of the porcine insulin-like growth factor binding protein-5 complementary deoxyribonucleic acid.
#accession JC4584
##molecule_type mRNA
##residues 1-271 ##label WHI
##cross-references GB:U41340
##experimental source skeletal muscle
COMMENT This protein has essential roles in the regulation and coordination of insulin-like growth factors action. This protein enhances the in vitro activity of the insulin-like growth factors, plays a role during myoblast proliferation and differentiation, and is important in the growth and development of muscle tissue.
KEYWORDS differentiation; growth factor; skeletal muscle
FEATURE
1-19 #domain signal sequence #status predicted #label SIG\
20-271 #product insulin-like growth factor binding protein-5 #status predicted #label MAT
SUMMARY #length 271 #molecular-weight 30323 #checksum 809
Query Match 4.7%; Score 128; DB 13; Length 271;
Best Local Similarity 37.5%; Pred. No. 8.02e-06;
Matches 30; Conservative 11; Mismatches 31; Indels 8; Gaps 7;
Db 2 vitavlllllaacagpqlgfvhcepcdekalamcpplgceelvkdpgccmcala 61
Qy 9 VVAVFVLLALCSRPVAVGQNSGFC-RCPEPAPRC-PA-GVSLVLD-GCGCCRCVAKQ 63

```

```

Db 62 egqscgvtyterc-aq-glrc 79
Qy 64 LGELCTE-RDPCDPHKLGLFC 82

RESULT 11
ENTRY JC1463 #type complete
TITLE insulin-like growth factor-binding protein 5 precursor - rat
ORGANISM #formal name Rattus norvegicus #common name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995
ACCESSIONS JC1463; A40403; F40403
REFERENCE JC1463
#authors Zhu, X.; Ling, N.; Shimaaki, S.
#journal Biochem. Biophys. Res. Commun. (1993) 190:1045-1052
#title Cloning of the rat insulin-like growth factor binding protein-5 gene and DNA sequence analysis of its promoter region.
#accession JC1463
##molecule_type DNA
##residues 1-271 ##label ZHU
##cross-references GB:L08275
REFERENCE A40403
#authors Shimaaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
#journal J. Biol. Chem. (1991) 266:10646-10653
#title Identification of five different insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human.
#cross-references MIM:91244847
#accession A40403
##molecule_type mRNA
##residues 1-271 ##label SHI
##cross-references GB:M62781
#accession F40403
##status preliminary
##molecule_type protein
##residues 20-25, 'X', 27-28, 'X', 30-36, 'X', 38-43, 'X', 45-51, 'X', 53, 'XX' ##label SH2
GENETICS
#introns 112/3; 188/3; 228/3
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE
1-19 #domain signal sequence #status predicted #label SIG\
20-271 #product insulin-like growth factor binding protein 5 #status predicted #label MAT\
191-262 #domain thyroglobulin type I repeat homology #label THY1
SUMMARY #length 271 #molecular-weight 30298 #checksum 1075
Query Match 4.6%; Score 125; DB 14; Length 271;
Best Local Similarity 36.0%; Pred. No. 2.33e-05;
Matches 27; Conservative 12; Mismatches 28; Indels 8; Gaps 7;
Db 7 llllaacavpqlgfvhcepcdekalamcpplgceelvkepgccmcalaegsc 66
Qy 14 VLLALCSRPVAVGQNSGFC-RCPEPAPRC-PA-GVSLVLD-GCGCCRCVAKQLGLFC 68
Db 67 gvyterc-aq-glrc 79
Qy 69 TE-RDPCDPHKLGLFC 82

```

```
RESULT 12
ENTRY   JP0076      #type complete
TITLE   nel protein - chicken
ORGANISM #formal name Gallus gallus #common name chicken
DATE    24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change
        01-Dec-1995
ACCESSIONS A38963; JP0076
REFERENCE  A38963
#authors   Matsuhashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
#journal   Taniguchi, S.; Hori, K.
#title     Dev. Dyn. (1995) 203:212-222
#accession A38963
#molecule_type mRNA
#residues 1-835 #label MAT
#cross-references DDBJ:D45365
#experimental_source 9-day embryo
REFERENCE  JP0076
#authors   Matsuhashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
#submission Taniguchi, S.; Hori, K.
#description A new gene, nel, encoding a Mr 93K protein with EGF-like
            repeats is strongly expressed in neural tissues of early
            stage chick embryos.
#accession JP0076
#molecule_type mRNA
#residues 1-835 #label MA2
#cross-references DDBJ:D45365
#experimental_source 9-day embryo
CLASSIFICATION #superfamily von Willebrand factor type C repeat homology
FEATURE
273-333    #domain von Willebrand factor type C repeat homology
            #label VMC
395-592    #region EGF-like repeats
SUMMARY    #length 835 #molecular-weight 93411 #checksum 7565
            4.6%; Score 126; DB 13; Length 835;
            Best Local Similarity 40.3%; Pred. No. 1.63e-05;
            Matches 25; Conservative 10; Mismatches 20; Indels 7; Gaps 6;
Db 274 ctmkgmttyrefeswtgdgck-nctcmngtvqcealicsdc-p-psalsay-vd--gkcc 327
Qy 103 CIFGTVYVRSGESFQSSCKYQCTCLDGAVGCMPL-CSDMVLRLSPDCPPRRVKLPKCC 161
Db 328 ke 329
Qy 162 EE 163

RESULT 13
ENTRY   B37252      #type complete
TITLE   insulin-like growth factor-binding protein 4 precursor -
        human
ALTERNATE_NAMES IGFBP-4; inhibitory insulin-like growth factor-binding
        protein; insulin-like growth factor-binding protein PC3
ORGANISM #formal name Homo sapiens #common name man
DATE    28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change
```

```
06-Sep-1996
ACCESSIONS B37252; B39842; A36549; A60712; A54650; A49801; A34419
REFERENCE  A37252
#authors   Shimaeaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.
#journal   Mol. Endocrinol. (1990) 4:1451-1458
#title     Molecular cloning of the cDNAs encoding a novel insulin-like
        growth factor-binding protein from rat and human.
#cross-references MUID:91133415
#accession B37252
#status    not compared with conceptual translation
#molecule_type mRNA
#residues 1-258 #label SHI
REFERENCE  A39842
#authors   Kiefer, M.C.; Mastarz, F.R.; Bauer, D.M.; Zapf, J.
#journal   J. Biol. Chem. (1991) 266:9043-9049
#title     Identification and molecular cloning of two new 30-kDa
        insulin-like growth factor binding proteins isolated from
        adult human serum.
#cross-references MUID:91225006
#accession B39842
#molecule_type mRNA
#residues 1-258 #label KIE
REFERENCE  A36549
#authors   Lafour, D.; Mohan, S.; Linkhart, T.A.; Baylink, D.J.; Strong,
        D.D.
#journal   Mol. Endocrinol. (1990) 4:1806-1814
#title     Inhibitory insulin-like growth factor-binding protein:
        cloning, complete sequence, and physiological regulation.
#cross-references MUID:91186988
#accession A36549
#molecule_type mRNA
#residues 1-50, 'A', '52-197', 'F', 199-258 #label LAT
REFERENCE  A60712
#authors   Perkel, V.S.; Mohan, S.; Baylink, D.J.; Linkhart, T.A.
#journal   J. Clin. Endocrinol. Metab. (1990) 71:533-535
#title     An inhibitory insulin-like growth factor binding protein
        (In-IGFBP) from human prostatic cell conditioned medium
        reveals N-terminal sequence identity with bone derived
        In-IGFBP.
#accession A60712
#molecule_type protein
#residues 22-26, 'X', 28-29, 'X', 31-35 #label PER
REFERENCE  A54650
#authors   Mohan, S.; Baylink, D.J.
#journal   Growth Regul. (1991) 1:110-118
#title     Evidence that the inhibition of TE85 human bone cell
        proliferation by agents which stimulate cAMP production may
        in part be mediated by changes in the IGF-II regulatory
        system.
#cross-references MUID:93091814
#accession A54650
#molecule_type protein
#residues 22-29, 'X', 31-42 #label MOH
#cross-references NCBIP:121076
#experimental_source TE85 osteosarcoma cells
#note      sequence extracted from NCB1 backbone
REFERENCE  A49801
#authors   Culouscou, J.M.; Shoyab, M.
#journal   Cancer Res. (1991) 51:2813-2819
#title     Purification of a colon cancer cell growth inhibitor and its
```


Best Local Similarity 36.9%; Pred. No. 5.32e-04;
Matches 24; Conservative 14; Mismatches 21; Indels 6; Gaps 6;

Db 1 mlpcalvaallitagrpelgde-aihcpcseeklarcrppvgceelvrepqgcaatc 59
| | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 6 MGFVRVAFVWLLALCSRPVAVGQNCSPCR-CPDEFAPRC-PA-GVS-LVLD-GGCGCRVC 60
Db 60 alglg 64
| | |
Qy 61 AKQLG 65

Search completed: Wed Sep 17 09:34:50 1997
Job time : 82 secs.

maryh@stic

4381-1

NeWSprinter20

Wed Sep 17 11:17:56 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

 M P S R L H
 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by IntelliGenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:32:15 1997; MapPar time 9.17 Seconds
 807.003 Million cell updates/sec
 Tabular output not generated.

Title: >DS-08-167-628-2
 Description: (1-349) from 3408040.ppe
 Perfect Score: 2713
 Sequence: 1 MTAASMGPRVAFVLLALC.....PGDNDFESLYRMYGDMA 349

Scoring table: PAM 150
 Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot34
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11

Statistics: Mean 44.133; Variance 66.349; scale 0.665

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2713	100.0	349	3	CTGF_HUMAN	CONNECTIVE TISSUE GRO 0.00e+00
2	2514	92.7	348	3	CTGF_MOUSE	CONNECTIVE TISSUE GRO 0.00e+00
3	1478	54.5	343	6	NOV_XENLA	NOV PROTEIN HOMOLOG P 0.00e+00
4	1460	53.8	351	6	NOV_CHICK	NOV PROTEIN PRECURSOR 0.00e+00
5	1450	53.4	353	6	NOV_COTJA	NOV PROTEIN PRECURSOR 0.00e+00
6	1415	52.2	357	6	NOV_HUMAN	NOV PROTEIN HOMOLOG P 0.00e+00
7	1152	42.5	375	2	CELO_CHICK	CEF-10 PROTEIN PRECUR 6.10e-277
8	1113	41.0	379	3	CYR6_MOUSE	CYR61 PROTEIN PRECURS 4.24e-266
9	133	4.9	254	5	IBP4_RAT	INSULIN-LIKE GROWTH F 5.18e-09

10	128	4.7	867	9	SSPO_BOVIN	SCO-SPONDIN (FRAGMENT	4.95e-08
11	125	4.6	271	5	IBP5_RAT	INSULIN-LIKE GROWTH F	1.88e-07
12	123	4.5	258	5	IBP4_HUMAN	INSULIN-LIKE GROWTH F	4.54e-07
13	121	4.5	455	1	60A_DROME	60A PROTEIN PRECURSOR	1.09e-06
14	116	4.3	254	5	IBP4_MOUSE	INSULIN-LIKE GROWTH F	9.35e-06
15	113	4.2	272	5	IBP1_RAT	INSULIN-LIKE GROWTH F	3.32e-05
16	113	4.2	281	10	VWF_HUMAN	VON WILLEBRAND FACTOR	3.32e-05
17	110	4.1	283	5	IBP1_BOVIN	INSULIN-LIKE GROWTH F	1.16e-04
18	111	4.1	271	5	IBP5_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
19	111	4.1	304	5	IBP2_RAT	INSULIN-LIKE GROWTH F	7.65e-05
20	111	4.1	305	5	IBP2_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
21	110	4.1	2274	6	MUC2_HUMAN	MUCIN 2 (INTESTINAL M	1.16e-04
22	109	4.0	272	5	IBP1_MOUSE	INSULIN-LIKE GROWTH F	1.75e-04
23	106	3.9	258	5	IBP4_BOVIN	INSULIN-LIKE GROWTH F	5.92e-04
24	105	3.9	272	5	IBP5_HUMAN	INSULIN-LIKE GROWTH F	8.86e-04
25	106	3.9	807	4	FSPO_RAT	F-SPONDIN PRECURSOR	5.92e-04
26	102	3.8	259	5	IBP1_HUMAN	INSULIN-LIKE GROWTH F	2.92e-03
27	102	3.8	231	5	IBP3_BOVIN	INSULIN-LIKE GROWTH F	2.92e-03
28	104	3.8	311	5	IBP2_CHICK	INSULIN-LIKE GROWTH F	1.32e-03
29	103	3.8	328	5	IBP2_HUMAN	INSULIN-LIKE GROWTH F	1.97e-03
30	103	3.8	500	10	VLI_HP6B	MAJOR CAPSID PROTEIN	1.97e-03
31	102	3.8	798	5	ITB7_HUMAN	INTEGRIN BETA-7 SUBUN	2.92e-03
32	104	3.8	313	5	HMCT_BOMMO	HEMOCYTIN PRECURSOR	1.32e-03
33	101	3.7	388	3	CSP_FLARE	CIRCUMSPOROZOITE PROT	4.32e-03
34	101	3.7	399	1	BM6A_MOUSE	BONE MORPHOGENETIC PR	4.32e-03
35	101	3.7	402	1	BMP8_HUMAN	BONE MORPHOGENETIC PR	4.32e-03
36	101	3.7	405	2	CSP_PLAFO	CIRCUMSPOROZOITE PROT	4.32e-03
37	101	3.7	412	2	CSP_PLAFT	CIRCUMSPOROZOITE PROT	4.32e-03
38	101	3.7	424	3	CSP_PLAFW	CIRCUMSPOROZOITE PROT	4.32e-03
39	100	3.7	442	3	CSP_PLAFW	CIRCUMSPOROZOITE PROT	6.38e-03
40	101	3.7	551	11	YD3_CAEEL	HYPOTHETICAL ZINC MET	4.32e-03
41	99	3.6	111	5	IBP5_BOVIN	INSULIN-LIKE GROWTH F	9.40e-03
42	99	3.6	291	5	IBP3_MOUSE	INSULIN-LIKE GROWTH F	9.40e-03
43	99	3.6	292	5	IBP3_RAT	INSULIN-LIKE GROWTH F	9.40e-03
44	98	3.6	806	5	ITB7_MOUSE	INTEGRIN BETA-7 SUBUN	1.38e-02
45	96	3.5	803	4	FSPO_XENLA	F-SPONDIN PRECURSOR	2.95e-02

ALIGNMENTS

RESULT	1	CTGF_HUMAN	STANDARD;	PRT;	349 AA.
ID	P29279;				
AC	01-DEC-1992 (REL. 24, CREATED)				
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.				
GN	CTGF.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;				
RX	MEDLINE; 91373462.				
RA	BRADHAM D.M., IGARASHI A., POTTER R.L., GROTEENDORST G.R.;				
RL	J. CELL BIOL. 114:1285-1294 (1991).				
RN	[2]				
RC	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;				

Listing for Mary Hale

Wed Sep 17 11:16:03 1997

MDLINE; 93187114.
RX IGARASHI A., BRADHAM D.M., OKOCHI H., GROTEENDORST G.R.;
RL J. DERMATOL. 19:642-643(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA OENAR B.S., WERNER A., YANG Z., GARNIER J.M., GENTZ R., LUESCHER T.F.;
RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
CC -!- SUBUNIT: MONOMER.
CC -!- ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE.
CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
CC -!- SIMILARITY: CONTAINS A VMFC DOMAIN.
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
DR EMBL; M92934; G180924; -.
DR EMBL; S56201; G266235; -.
DR EMBL; X78947; G474934; -.
DR PIR; A40551; A40551.
DR PIR; S44205; S44205.
DR MIM; 121009; -.
DR PROSITE; PS00222; IGF BINDING.
DR GROWTH FACTOR BINDING; SIGNAL; ALTERNATIVE SPLICING.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 101 167 VMFC.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 284 323 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
FT CARBOHYD 28 28 POTENTIAL.
FT CARBOHYD 225 225 POTENTIAL.
FT VARSPLIC 172 198 MISSING (IN SHORT FORM).
SQ SEQUENCE 349 AA; 38069 MW; C21E9662 CRC32;
Query Match 100.0%; Score 2713; DB 3; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasmgprvrvafvllalcerpavqncgpcrcpdepaprcpagsvlvldgcgcrcvc 60
Qy 1 MTAASMGPRVRVAFVLLALCSRPAVGQNCSPCRCPDEPAPRCPAGSVLVDGCGCRVC 60
Db 61 akqlgelcterdpdphkglfcdifgsparnkigvctakdgcpcifggtvyrsgesfqsac 120
Qy 61 AKQLGELCTERDPCDPHKGFLCDFGSPANKIGVCTAKDGCPCIFGGTVYRSGESFQSSC 120
Db 121 kyqctcldgavcmplcamdvrlspdcppfrvklpgkceewvcdpdkdtvvpapaa 180
Qy 121 KYQCTCLDGA VCMPLCAMDVRLSPDCPPFRV KLPKCEEWVCDPDKDTVVPAPAA 180
Db 181 avrledtfgpdpmtmrancivqtetwaacsctcmgtstvtndnaacrlckgsrlcmvr 240
Qy 181 AYRLDFTFGPDPMTMRANCLIVQTETWASCSCTCMGTSTVTRVNDNASCRLEKOSRLCMVR 240
Db 241 pceadleenikkkgkckirtpkiskpikfelsgctamktyrakfcgvcctdgrctphrttt 300

Listing for Mary Hale

Wed Sep 17 11:16:03 1997

MDLINE; 93187114.
RX IGARASHI A., BRADHAM D.M., OKOCHI H., GROTEENDORST G.R.;
RL J. DERMATOL. 19:642-643(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA OENAR B.S., WERNER A., YANG Z., GARNIER J.M., GENTZ R., LUESCHER T.F.;
RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
CC -!- SUBUNIT: MONOMER.
CC -!- ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE.
CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
CC -!- SIMILARITY: CONTAINS A VMFC DOMAIN.
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
DR EMBL; M92934; G180924; -.
DR EMBL; S56201; G266235; -.
DR EMBL; X78947; G474934; -.
DR PIR; A40551; A40551.
DR PIR; S44205; S44205.
DR MIM; 121009; -.
DR PROSITE; PS00222; IGF BINDING.
DR GROWTH FACTOR BINDING; SIGNAL; ALTERNATIVE SPLICING.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 101 167 VMFC.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 284 323 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
FT CARBOHYD 28 28 POTENTIAL.
FT CARBOHYD 225 225 POTENTIAL.
FT VARSPLIC 172 198 MISSING (IN SHORT FORM).
SQ SEQUENCE 349 AA; 38069 MW; C21E9662 CRC32;
Query Match 100.0%; Score 2713; DB 3; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasmgprvrvafvllalcerpavqncgpcrcpdepaprcpagsvlvldgcgcrcvc 60
Qy 1 MTAASMGPRVRVAFVLLALCSRPAVGQNCSPCRCPDEPAPRCPAGSVLVDGCGCRVC 60
Db 61 akqlgelcterdpdphkglfcdifgsparnkigvctakdgcpcifggtvyrsgesfqsac 120
Qy 61 AKQLGELCTERDPCDPHKGFLCDFGSPANKIGVCTAKDGCPCIFGGTVYRSGESFQSSC 120
Db 121 kyqctcldgavcmplcamdvrlspdcppfrvklpgkceewvcdpdkdtvvpapaa 180
Qy 121 KYQCTCLDGA VCMPLCAMDVRLSPDCPPFRV KLPKCEEWVCDPDKDTVVPAPAA 180
Db 181 avrledtfgpdpmtmrancivqtetwaacsctcmgtstvtndnaacrlckgsrlcmvr 240
Qy 181 AYRLDFTFGPDPMTMRANCLIVQTETWASCSCTCMGTSTVTRVNDNASCRLEKOSRLCMVR 240
Db 241 pceadleenikkkgkckirtpkiskpikfelsgctamktyrakfcgvcctdgrctphrttt 300

QY 1 MTAASMGPRVAVFVLLALCSRPVAVGQNCSPCRCPDEPAPRCPAGVSIIVLDGCGCCRV 60
 Db 60 akqlgelctetdpcdkhglfcdgspgpanrkigvctakdgapcvfgsvyrgesfgsac 119
 QY 61 AKQGLGELCTERDPCDPHKGFLCDFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSC 120
 Db 120 kyqctcldgavcgvpcldmrvlpdpdpfprvklpgkckewwcdedkdrtaavpala 179
 QY 121 KYQCTCLDGA VGCMP LCMNDVRLSPDCPFPRVRKLPKCCCEWVCDPEKDTQVWGPALA 180
 Db 180 avrledtfgpdpmmranclvqtewesactcmglsrtvndntfclerkasrlcmvr 239
 QY 181 AYRLDFTGPDPTMIRANCLVQTTWESACSKTCGMGISTRTVNDNASCRLEKQSRLOMR 240
 Db 240 pceadleenikkkkciirtpkiakpvkfclagctsvtktyakfgvtdgrcctphrttt 299
 QY 241 PCEADLEENIKKKKCIIRTPKISPKIFELSGCTSMKTYRAKTCGVCTDGRCTPHRTTT 300
 Db 300 lpvefkpdpgeimkknmmfiktacchyncpndndifaelvyrkmygdma 348
 QY 301 LPVEFKPDPGEVMMKNMMFIKTACHYNCPCGNDIFESLYRYKMYGDMA 349

RESULT 3
 ID NOV XENIA STANDARD; PRT; 343 AA.
 AC P51609;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NOV PROTEIN HOMOLOG PRECURSOR (XNOV).
 GN XNOV.
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA YING Z., KING M.L.;
 RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; U37063; G1176415; -.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 18
 FT CHAIN 19 343
 FT DOMAIN 93 159
 FT DOMAIN 249 323
 FT DISULFID 249 286
 FT DISULFID 266 300
 FT DISULFID 277 316
 FT DISULFID 280 318
 FT DISULFID 285 322
 FT CARBOHYD 265 265
 SQ SEQUENCE 343 AA; 38070 MW; BDF3BCA4 CRC32;

Query Match 54.5%; Score 1478; DB 6; Length 343;
 Best Local Similarity 53.7%; Pred. No. 0.00e+00;
 Matches 180; Conservative 69; Mismatches 78; Indels 8; Gaps 8;

Db 1 mtp-hlalefil-l-iqqvasqkpsqcdqcppeep-pacapavllilldgcccpcarqe 56
 QY 6 MGEVRFVAVFVLLALCSRPVAVGQNCSPCRCPDEPAPRCPAGVSIIVLDGCGCCRVCAKQL 64
 Db 57 geacahlnpcqcdkglkycefnadpmetgtcmaleqncavfdgvyvnrresfgpckvyc 116
 QY 65 GELCTERDPCDPHKGFLCDFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQC 124
 Db 117 tclnghgvcvprcnldlllpgdpdpfprvklpgkckewwcdedkdrtaavpala 175
 QY 125 TCLDGA VGCMP LCMNDVRLSPDCPFPRVRKLPKCCCEWVCDPEKDTQVWGP-ALAYR 183
 Db 176 peatlgidaedtsafaciqtewesactcmglsrtvndntfclerkasrlcmvr 235
 QY 184 LEDTFGPDPTMIRANCLVQTTWESACSKTCGMGISTRTVNDNASCRLEKQSRLOMR 240
 Db 236 eegghvkekkgkcvrvrttkpihfykntsvqpkfkgcgsgdgrcctphrttt 295
 QY 244 ADLEENI-KRGGKCIIRTPKISPKIFELSGCTSMKTYRAKTCGVCTDGRCTPHRTTLP 302
 Db 296 vefvcpdkriivkpmvviatcchyncpndndifaelvyrkmygdma 330
 QY 303 VEFVCPDPGEVMMKNMMFIKTACHYNCPCGNDIFE 337

RESULT 4
 ID NOV CHICK STANDARD; PRT; 351 AA.
 AC P28586;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NOV PROTEIN PRECURSOR.
 GN NOV.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BROWN LEHORN;
 RX MEDLINE; 92107157.
 RA MALOISEL V., MARTINERIE C., DAMBRINE G., PLASSIART G., BRISAC M.,
 RA CROCHET J., PERBAL B.;
 RL MOL. CELL. BIOL. 12:10-21(1992).
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
 CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION
 CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -1- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 CC ADULT KIDNEY.
 CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
 CC MUSCLE AND INTESTINE, IN THE EMBRYO. LONG AND LESS SO IN BRAIN AND
 CC SPLEEN, IN ADULT CHICKEN.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; X59284; G63703; -.

[illegible]

RESULT	5	NOV	COTUJA	STANDARD;	PRT;	353	AA.
ID	P42642;						
AC	01-NOV-1995	(REL. 32,	CREATED)				
DDT	01-NOV-1995	(REL. 32,	LAST SEQUENCE UPDATE)				
DDT	01-OCT-1996	(REL. 34,	LAST ANNOTATION UPDATE)				
DDT	NOV	PROTEIN	PRECURSOR.				
DE	NOV	COTURNIX	JAPONICA (JAPANESE QUAIL).				
GN	COTURNIX	JAPONICA	(JAPANESE QUAIL).				
OS	EUKARYOTA;	METAZOA;	CHORDATA;	VERTEBRATA;	TETRAPODA;	AVES;	NEOGNATHAE;
OC	GALLIFORMES.						
CC	{1}						
CC	SEQUENCE FROM N.A.						
RN	WELSKIRCHEN R.,	BISTER K.;					
R2	SUBMITTED (AUG-1994)	TO EMBL/GENBANK/DBJ	DATA BANKS.				
R2	-1- FUNCTION:	IMMEDIATE-EARLY	PROTEIN LIKELY TO PLAY A	ROLE IN CELL			
CCC							

[illegible]

RESULT	6	STANDARD;	PRT;	357 AA.
NOV HUMAN				
AD P48745;				
IC				
DDT	01-FEB-1996 (REL. 33, CREATED)			
DDT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)			
DDT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DDT	NOV PROTEIN HOMOLOG PRECURSOR (NOVH).			
NOV				
CGN				
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPO			
OC	EUTHERIA; PRIMATES.			

		245	rvtndnpdcklketricevrvpcgpgsyasalkkgkctktkkpsvrvcyagcsavky	304
Db			: : : : : : : : : :	
		279	rrtvdnasclrekgosrlcmwrvpceadleenikkkgkcirtppkiSKPIKfELSGCTSMKIT	279
Qy			: : : : : : : : : :	
		305	tpkvsgscvgdrctpqatrtvkirfrddgetfkavmmiqacrcnycphaneay-p-	362
Db			: : : : : : : : : : :	
		280	fakfcgvctdgroccrthrtttlefvfePKCPDGEVMMKNMFIKTCACHYNCPGDNDFEFL	339
Qy			: : : : : : : : : :	
		363	fyv 365	
Db			:	
		340	yvr 342	
Qy			:	
		RESULT	8	
ID	CYR6 MOUSE	STANDARD;	PRT;	379 AA.
AC	P18406;			
DT	01-NOV-1990 (REL. 16, CREATED)			
DD	01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)			
DE	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	CYR61 PROTEIN PRECURSOR (3CH61).			
GN	CYR61			
OS	MUS MUSCULUS (MOUSE).			
OOC	EUKARYOTA; METAZOA;			
OCC	EUTHERIA; RODENTIA.			
	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=BALB/C / 3T3;			
RC	MEDLINE; 90287146.			
RX	O'BRIEN T.P., YANG G.P., SANDERS L., LAU L.F.;			
RA	MOL. CELL. BIOL. 10:3569-3577(1990).			

[2]	
RN	SEQUENCE FROM N.A.
RP	STRAIN-AJ; TISSUE=EMBRYONIC FIBROBLAST;
RC	MEDLINE; 91288203.
RX	LATINIC B.V., O'BRIEN T.P., LAU L.F.;
RL	NUCLEIC ACIDS RES. 19:3261-3267(1991).
CC	-!- FUNCTION: CYR61 MIGHT ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING PROTEINS.
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
CC	-!- INDUCTION: BY GROWTH FACTORS.
CC	-!- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN, AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST IN LUNG.
CC	-!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
CC	-!- SIMILARITY: CONTAINS A WVFC DOMAIN.
CC	-!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
EMBL	EMBL; M32490; G309206; -.
DR	EMBL; X56790; G50633; -.
DR	PIR; A35669; A35669.
DR	PROSITE; PS00222; IGF BINDING.
DR	GROWTH FACTOR BINDING; SIGNAL.
KW	SIGNAL 1 24 POTENTIAL.
FT	CHAIN 25 379 CYR61 PROTEIN.
FT	DOMAIN 98 164 WVFC.
FT	DOMAIN 284 358 CTCK.
FT	DISULFID 284 321 BY SIMILARITY.
FT	DISULFID 301 335 BY SIMILARITY.
FT	DISULFID 312 351 BY SIMILARITY.

[illegible]

RESULT	9	STANDARD;	PRT;	254 AA.
IPB4 RAT				
AD	P21744;			
IC	P21744;			
DDT	01-MAY-1991 (REL. 18, CREATED)			
DDT	01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)			
DDT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)			
DDT	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)			
DE	(IBP-4) (IGF-BINDING PROTEIN 4).			
DE	IGFBP4 OR IGFBP-4.			
OS	RATTUS NORVEGICUS (RAT).			
OS	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
OC	[1]			
IRN	SEQUENCE FROM N.A.			
RP	MEDLINE; 91133415.			
RX	SHIMASAKI S., UCHIYAMA F., SHIMONAKA M., LING N.;			
RA	MOL. ENDOCRINOL. 4:1451-1458(1990).			
RL	[2]			
IRN	SEQUENCE FROM N.A.			
RP	STRAIN=SPRAGUE-DAWLEY;			
RC	MEDLINE; 93176147.			
RX	GAO L., LING N., SHIMASAKI S.;			
RA	BIOCHEM. BIOPHYS. RES. COMMUN. 190:1053-1059(1993).			
RL				

RN SEQUENCE OF 22-61.
 RC TISSUE=SERUM;
 RA SHIMONAKA M., SCHROEDER R., SHIMASAKI S., LING N.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 165:189-195(1989).
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC INTERACTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC PROMOTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -1- BINDS IGF-II MORE THAN IGF-I.
 CC -1- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
 DR EMBL; L08276; E73337; -.
 DR PIR; B33570; B33570.
 DR PIR; A37252; A37252.
 DR PIR; J01464; J01464.
 DR PIR; E40403; E40403.
 DR PROSITE; PS00222; IGF BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN 1.
 KW GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 21
 FT CHAIN 22 254 INSULIN-LIKE GROWTH FACTOR BINDING
 FT CARBOHYD 125 125 PROTEIN 4.
 FT DOMAIN 196 245 POTENTIAL.
 FT SEQUENCE 234 AA; 27745 MW; 3B5316DC CRC32;
 Query Match 4.9%; Score 133; DB 5; Length 254;
 Best Local Similarity 39.7%; Pred. No. 5,18e-09;
 Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
 Db 1 mlpfglaallaaqprslqde-aihcpcseeklarcpvpgvgeelvrepqgcgcac 59
 QY 6 MGPVRVAVVLLALCSRAVQNGSGPCR-CPDEAPRGC-PA-GVS-LVLD-GCGGCRVC 60
 Db 60 algiqmpc 67
 QY 61 AKQLGELC 68
 RESULT 10
 ID SSPO BOVIN STANDARD; PRT; 867 AA.
 AC P98167;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DE SCO-SPONDIN (FRAGMENT).
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; ANIMALIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA;
 OC THERIA; EUTHERIA; ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EPIDYMOCTE;
 RX MEDLINE; 96338614.
 RA GORON S., MONNERIE H., MEINIEL R., CREVEAUX I., LEHMANN W.,
 RA LAMALLE D., DASTUGUE B., MEINIEL A.;
 RL J. CELL SCI. 109:1053-1061(1996).
 CC -1- FUNCTION: INVOLVED IN THE MODULATION OF NEURONAL AGGREGATION.
 CC -1- TISSUE SPECIFICITY: SUBCELLULAR ORGAN.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
 CC -1- SIMILARITY: CONTAINS AT LEAST 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS AT LEAST 3 LDL-RECEPTOR CLASS A DOMAINS.
 DR EMBL; X93922; E213357; -.
 KW GLYCOPROTEIN; CELL ADHESION; CALCIUM-BINDING; REPEAT; EGF-LIKE DOMAIN.
 FT NON TER 1 142
 FT DOMAIN 103 142 EGF-LIKE.
 FT DOMAIN 143 180 EGF-LIKE.
 FT DOMAIN 506 544 LDL-RECEPTOR CLASS A.
 FT DOMAIN 663 701 LDL-RECEPTOR CLASS A.
 FT DOMAIN 723 761 LDL-RECEPTOR CLASS A.
 FT CARBOHYD 88 88 POTENTIAL.
 FT CARBOHYD 309 309 POTENTIAL.
 FT CARBOHYD 409 409 POTENTIAL.
 FT NON TER 867 867
 SQ SEQUENCE 867 AA; 91817 MW; B1224081 CRC32;
 Query Match 4.7%; Score 128; DB 9; Length 867;
 Best Local Similarity 28.3%; Pred. No. 4.95e-08;
 Matches 32; Conservative 21; Mismatches 48; Indels 12; Gaps 12;
 Db 81 cpxdgcgvntcagelyvhacvpcplctcdidsgatcpdr-pcggp-gcwcpgagvlgag 138
 QY 35 CPDEAPRCPAGVSLVLDGCGCRV-CAKQLGEL-CT-ERPCDPHGKGLFCDFGSPANK 91
 Db 139 -gcvvprq-cpcldvgdrywpqrvktcdq1-cvcqdgprprcqsdcavn 188
 QY 92 IGVCT-AKDGAPCIFGTVYRSGESFOSCKYQCTCLDG-AVGCMF-L-CSMD 140
 RESULT 11
 ID IBP5 RAT STANDARD; PRT; 271 AA.
 AC P24594;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP-5)
 DE (IBP-5) (IGF-BINDING PROTEIN 5).
 GN IGFBP5 OR IGFBP-5.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-53.
 RC TISSUE=OVARY;
 RX MEDLINE; 91244847.
 RA SHIMASAKI S., SHIMONAKA M., ZHANG H.-P., LING N.;
 RL J. BIOL. CHEM. 266:10646-10653(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE; 93176146.
 RA ZHU X., LING N., SHIMASAKI S.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:1045-1052(1993).
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -1- TISSUE SPECIFICITY: MOSTLY IN KIDNEY.


```

RX MEDLINE; 91141487.
RA MOHN K.L., MELBY A.E., TEWARI D.S., LAZ T.M., TAUB R.A.;
RL MOL. CELL. BIOL. 11:1393-1401(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93149132.
RA OOI G.T., TSENG L.Y.H., TRAN M.O., RECHLER M.M.;
RL MOL. ENDOCRINOL. 6:2219-2228(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE; 94250701.
RA LACSON R.G., OEHLER D., YANG E., GOSWAMI R., UNTERMAN T.G.;
RL BIOCHIM. BIOPHYS. ACTA 1218:95-98(1994).
RN [5]
RP SEQUENCE OF 26-59.
RX MEDLINE; 90322923.
RA UNTERMAN T.G., OEHLER D.T., GOTWAY M.B., MORRIS P.W.;
RL ENDOCRINOLOGY 127:789-797(1990).
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -!- BINDS EQUALLY WELL IGF-I AND IGF-II.
CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR EMBL; M58634; G204733; -.
DR EMBL; M89791; G204737; -.
DR EMBL; L22979; G1098473; -.
DR FIR; A36082; A36082.
DR FIR; A37398; A37398.
DR FIR; A39683; A39683.
DR PROSITE; PS00222; IGF BINDING.
DR PROSITE; PS00484; THYROGLOBULIN_1.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 272
FT INSULIN-LIKE GROWTH FACTOR BINDING
FT PROTEIN 1.
FT DOMAIN 215 264 THYROGLOBULIN TYPE I.
FT SITE 259 261 CELL ATTACHMENT SITE.
FT CONFLICT 79 79 R -> A (IN REF. 1 AND 4).
FT CONFLICT 111 112 AA -> PP (IN REF. 1).
FT CONFLICT 201 201 A -> R (IN REF. 3).
FT CONFLICT 265 265 H -> N (IN REF. 1).
SQ SEQUENCE 272 AA; 29684 MW; 59EDF790 CRC32;

```

```

Query Match 4.2%; Score 113; DB 5; Length 272;
Best Local Similarity 42.9%; Pred.No. 3.32e-05;
Matches 21; Conservative 4; Mismatches 18; Indels 6; Gaps 6;

Db 43 cpvpaas-cpe-isrp-agcgccptcalplgaacgvatarc-aq-glsc 86
QY 35 CPDEPAPRCFAGVSLVLDGCGCCRCVCAKQLGELC-TERDPCDPHKLFC 82

```

Search completed: Wed Sep 17 09:33:10 1997
Job time : 55 secs.

maryh@stic

4382-1

NeWSprinter20

Wed Sep 17 11:20:33 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPORH

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:30:25 1997; Maspar time 7.73 seconds
Tabular output not generated. 516.210 Million cell updates/sec

Title: >US-08-386-680-2
Description: (1-349) from 5985270.pep
Perfect Score: 2713
Sequence: 1 MTAASMGPRVAVFWLLALC.....PGNDIFESLYYRKMYGDMA 349

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.775; Variance 141.735; scale 0.231

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	2713	100.0	349	20	Connective tissue gro	3.93e-253
2	2713	100.0	349	16	Connective tissue gro	3.93e-253
3	2521	92.9	348	5	Beta-IG-M2.	4.65e-234
4	1460	53.8	351	6	Chicken nov protein.	4.62e-129
5	1113	41.0	379	5	Beta-IG-M1.	5.50e-95
6	1002	36.9	375	16	Connective tissue gro	3.88e-84
7	644	23.7	205	6	Homologous to chicken	1.93e-49

8	579	21.3	72	6	R31610	Fragment XXVI with ho	3.25e-43
9	430	15.8	84	6	R31602	Chicken nov protein f	4.30e-29
10	412	15.2	71	6	R31609	Encoded by chicken no	2.09e-27
11	405	14.9	84	6	R31603	Polypeptide X homolog	9.43e-27
12	395	14.6	124	10	R46078	CYR61 like protein.	8.10e-26
13	314	11.6	72	6	R31605	Homologous to chicken	2.49e-18
14	312	11.5	70	6	R31604	Chicken nov protein f	3.79e-18
15	278	10.2	75	6	R31601	Chicken nov protein f	4.66e-15
16	268	9.9	76	6	R31600	Chicken nov protein f	3.72e-14
17	197	5.8	22	6	R31612	Fragment XXXI homolo	1.68e-04
18	138	5.1	22	6	R31611	Fragment XXX encoded	5.99e-03
19	125	4.6	271	5	R26994	Rat IGFBP-5.	6.48e-02
20	123	4.5	184	19	R98994	Vascular IBP-like gro	9.29e-02
21	123	4.5	258	4	R22253	Sequence of insulin-1	9.29e-02
22	123	4.5	258	4	R21688	Sequence of insulin-1	9.29e-02
23	121	4.5	455	19	W00231	Drosophila morphogeni	1.33e-01
24	121	4.5	455	9	R47261	Pre-pro 60A.	1.33e-01
25	121	4.5	455	10	R60966	Drosophila 60A morpho	1.33e-01
26	121	4.5	455	10	R57981	Drosophila 60A morpho	1.33e-01
27	121	4.5	455	7	R33413	Drosophila 60A morpho	1.33e-01
28	121	4.5	455	7	R33935	Drosophila 60A full le	1.33e-01
29	121	4.5	455	9	R47298	60A.	1.33e-01
30	121	4.5	455	10	R50208	Drosophila 60A morpho	1.33e-01
31	121	4.5	455	10	R60962	Drosophila 60A morpho	1.33e-01
32	121	4.5	455	9	R46745	Drosophila sp. 60(A)	1.33e-01
33	117	4.3	282	13	R79101	Prostaglandin 12 (PGI	2.72e-01
34	117	4.3	282	13	R79102	Prostaglandin 12 (PGI	2.72e-01
35	113	4.2	193	3	P60463	Sequence of C-terminu	5.50e-01
36	113	4.2	2813	3	P60053	Sequence of human von	5.50e-01
37	112	4.1	2813	3	P60053	Sequence of von Wille	6.56e-01
38	109	4.0	1218	19	W05833	Human Serrate-1 (HSL	1.11e+00
39	105	3.9	272	19	R95329	Insulin-like growth f	2.21e+00
40	105	3.9	272	5	R26995	Human IGFBP-5.	2.21e+00
41	105	3.9	272	5	R25700	IGFBP6.	2.21e+00
42	105	3.9	272	10	R55084	Human insulin-like gr	2.21e+00
43	106	3.9	807	8	R44241	E-spondin (EP5-9).	1.86e+00
44	103	3.8	328	1	R04908	EcoRI-EcoRI fragment	3.11e+00
45	103	3.8	532	17	R83016	Recombinant papilloma	3.11e+00

ALIGNMENTS

RESULT	1
ID	W11302 standard; Protein; 349 AA.
AC	W11302;
DT	18-MAR-1997 (first entry)
DE	Connective tissue growth factor.
KW	Connective tissue growth factor; CTGF; human; connective tissue cell;
KW	Proliferative disease; platelet-derived growth factor; PDGF; development;
KW	tissue growth; repair; umbilical vein endothelial cell; HUVE cell;
KW	antibody; wound healing; cancer; fibrotic disease; atherosclerosis;
KW	inhibitor; protease degradation; growth factor; therapy.
OS	Homo sapiens.
PN	US5585270-A.
PD	17-DEC-1996.
PF	30-AUG-1991; 752427.
PR	30-AUG-1991; US-752427.
PR	14-DEC-1993; US-167628.
PR	10-FEB-1995; US-386680.
PA	(USF-) UNIV SOUTH FLORIDA.

PI Bradham DM, Grotendorst GR;
DR WPI; 97-051180/05.
PT New nucleic acid encoding connective tissue growth factor - useful
PT for accelerating wound healing, also for diagnosis and treatment of
PT proliferative disease
PS Claim 9; Column 15-18; lipp; English.
CC This sequence represents the human connective tissue growth factor
CC (CTGF). CTGF is related immunologically and biologically to
CC platelet-derived growth factor (PDGF), but is encoded by an unrelated
CC gene. CTGF is thought to play a significant role in the normal
CC development, growth, and repair of human tissue, similarly to PDGF. The
CC cDNA encoding this sequence was isolated by screening a cDNA library from
CC human umbilical vein endothelial (HUVE) cells with anti-PDGF antibodies.
CC CTGF can be used to accelerate wound healing. Also, elevated levels of
CC CTGF may be diagnostic of proliferative diseases involving outgrowth of
CC connective tissue cells, such as cancer, fibrotic disease and
CC atherosclerosis. All of these diseases can be treated with reagents
CC reactive with CTGF, such as antibodies (which can also serve as assay
CC reagents). Antisense nucleic acids, and ribozymes could also be used to
CC inhibit CTGF production. The advantage with using CTGF is that it is
CC more stable, and less susceptible to protease degradation than PDGF, and
CC other growth factors involved in wound healing. This is believed to be
CC due to the high Cys content.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 20; Length 349;
Best Local Similarity 100.0%; Pred. No. 3,93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasgprvrvafvllalcarpavqncsgpcrcpdepaprcpagvalvldgcgcrcvc 60
QY 1 MTAASMGPRVRVAFVLLALCSREAVGQNCSPCRCPDEPAPRCPCAGVSLVLDGCGCCRCVC 60
Db 61 akqlgelcterdpchphkglfcdgspanrki:gvctakdgapcifggtvyrsgesfqscc 120
QY 61 AKQLGELCTERDPCDPHKGFLCDFGSPANRKIGVCTAKDGAFCIFGTVYRSGESFQSSC 120
Db 121 kyqctcldgavcmplcsmdvrlpdpdpfprrvklpgkceewcdepkdtvvgpala 180
QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSPDCPFPRRVKLP GKCEEWCEDEPKDQTVVGPALA 180
Db 181 avrledtfpgdptmirancivqtewasactcmgmistrvtnndnasrleqkqrlcmvr 240
QY 181 AVRLEDTFGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNNDNASCRLEKQSRLCMVR 240
Db 241 pceadleenikkkgkciirtpkiskpkiflegctamkyrakfcgvctdgrcctphrttt 300
QY 241 PCEADLEENIKKGKCIIRTPKISKPKIFELSGCTSMKYRAKFCGVCTDGRCTPHRTTT 300
Db 301 lpvefkcpdgevmknmmfiktacachyncpgndnifealyrkmvgdma 349
QY 301 LPVEFKCPDGEVMKNMMFIKTCACHYNCPGNDNIFEALYRKMVGDMA 349

RESULT 2
ID R79964 standard; Protein; 349 AA.
AC R79964;
DT 12-JUN-1996 (first entry)
DE Connective tissue growth factor.

KW Connective tissue growth factor; CTGF; wound healing; vulnery;
KW cell proliferation; cancer; fibrosis; atherosclerosis; diagnosis
OS Homo sapiens.

FH Key Location/Qualifiers
FT Modified site 28
FT /label= N-glycosylation_site
FT Modified site 225
FT /label= N-glycosylation_site
PN US5408040-A.
PD 18-APR-1995; 752427.
PF 30-AUG-1991; 752427.
PR 30-AUG-1991; US-752427.
PR 14-DEC-1993; US-167628.
PA (USF-) UNIV SOUTH FLORIDA.
PI Bradham DM, Grotendorst GR;
PI WPI; 95-161147/21.
DR N-PSDB; T04226.

DR New connective tissue growth factor - used to develop prods. for
PT wound healing and for diagnosis and therapy of cell proliferative
PT disorders.

PS Claim 1; Column 19-20; l2pp; English.
CC Novel human connective tissue growth factor (CTGF) (R79964)
CC is related immunologically and biologically to platelet-derived
CC growth factor (PDGF), but is the product of a distinct gene.
CC CTGF is mitogenic and also a chemotactic agent for cells. It is
CC produced by endothelial and fibroblastic cells, and probably acts
CC as a growth factor in wound healing. Recombinant CTGF can be obt.
CC by expression of cDNA clone DB60R32 (T04226) in transformed host
CC cells. It is used to accelerate wound healing, and to raise
CC antibodies useful in detecting disorders associated with overgrowth
CC of cells, such as cancer, fibrotic diseases and atherosclerosis.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 16; Length 349;
Best Local Similarity 100.0%; Pred. No. 3,93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasgprvrvafvllalcarpavqncsgpcrcpdepaprcpagvalvldgcgcrcvc 60
QY 1 MTAASMGPRVRVAFVLLALCSREAVGQNCSPCRCPDEPAPRCPCAGVSLVLDGCGCCRCVC 60
Db 61 akqlgelcterdpchphkglfcdgspanrki:gvctakdgapcifggtvyrsgesfqscc 120
QY 61 AKQLGELCTERDPCDPHKGFLCDFGSPANRKIGVCTAKDGAFCIFGTVYRSGESFQSSC 120
Db 121 kyqctcldgavcmplcsmdvrlpdpdpfprrvklpgkceewcdepkdtvvgpala 180
QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSPDCPFPRRVKLP GKCEEWCEDEPKDQTVVGPALA 180
Db 181 avrledtfpgdptmirancivqtewasactcmgmistrvtnndnasrleqkqrlcmvr 240
QY 181 AVRLEDTFGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNNDNASCRLEKQSRLCMVR 240
Db 241 pceadleenikkkgkciirtpkiskpkiflegctamkyrakfcgvctdgrcctphrttt 300
QY 241 PCEADLEENIKKGKCIIRTPKISKPKIFELSGCTSMKYRAKFCGVCTDGRCTPHRTTT 300
Db 301 lpvefkcpdgevmknmmfiktacachyncpgndnifealyrkmvgdma 349
QY 301 LPVEFKCPDGEVMKNMMFIKTCACHYNCPGNDNIFEALYRKMVGDMA 349

QY 301 LPVEFKCPDGEVMKKNMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349

RESULT 3

ID R25566 standard; Protein; 348 AA.

AC R25566;

CD 18-JAN-1993 (first entry)

DE Beta-IG-M2.

DD Transforming growth factor beta; induced; CEF-10; v-src; chicken;

EW embryo; fibroblasts; TGF-beta.

KW Mus musculus.

OS EP-495674-A.

PN 22-JUL-1992.

PD 17-JAN-1992; 300429.

PR 18-JAN-1991; US-642991.

PP 10-JAN-1992; US-816270.

RA (BRIM) BRISTOL-MYERS SQUIBB CO.

RB Brunner AM, Chinn J, Neubauer MG, Purchio AF;

RI WPI; 92-243508/30.

RS NPDSB; Q26422.

RT TGF-beta induced gene family - encodes proteins involved in growth and differentiation effects of TGF-beta-1

SC Claim 3; Fig 2; 35pp; English.

SO The protein sequence was deduced from the DNA sequence obtd. by screening a cDNA library made from AKR-2B mouse cells induced with TGF-beta1 and cyclohexamide with two probes with cyclohexamide and TGF-beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1. Beta-IG-M2 displays 50 percent homology to the CEF-10 protein induced by v-src in chicken embryo fibroblasts. Residues 52-59 of beta-IG-M2 conform to the GGCXCXC motif reported in the amino half of insulin-like growth factor (IGF) binding proteins. The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contains an amino acid sequence with strong homology to a motif found near C-terminal of the malarial circumsporozoite (CS) protein, which is highly conserved among all species of malarial parasites sequenced to date (designated region II). This motif is also found in other proteins which have cell adhesive properties that mediate cell-cell and cell-extracellular matrix interactions, such as progerin, thrombospondin, and TRAP. The proteins encoded by TGF-beta induced genes are likely to be involved in mediation of the biological effects of TGF-beta relating to cell growth and differentiation. See also R25565.

Sequence 348 AA.

```
Query Match      92.9%; Score 2521; DB 5; Length 348;
Best Local Similarity 91.1%; Pred. No. 4,65e-234;
Matches 318; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

bb      1 mlasvagi.lsl.v-llalctrptgdcagcqaacaaahphcpagvislvdgcgcrcv 59
        ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
yy      1 MTAASGPVRVAFVLLALCSRPAGVQNCSPCRCPDEPAPRCPPAGVISLVDGCGCRCVC 60

bb      60 akqlgelcterdpdcbphkglfcdfgspnrrkigvctakdgapcvfsgsvyrgesfqssc 119
        ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
yy      61 AKQLGELCTERDPCDPHKGFLCDFGSPNRRKIGVCTAKDGAPCIFGGTVYRSGESFQSSC 120

bb      120 kyqctldgavgcvp.lcamdvrlpapdcppfrvrklpgkcceewcwdepkdxrtavgpala 179
        ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

121	KYQCTCLDGA VGCMPLCSMDVRLPSDFPRFRVKLFGKCEEWCDEPKDQTWGPALA	180
QY		
180	ayrledtfqdpmmranclvqtswsacsktcgmgiastvtndnfcrlkqgrlcmvr	239
Db		
181	AYRLEDTFGDPDMIRANCLVQTTSWSACSKTCGMGISTRVNDNASCRIEKGSRLCWR	240
QY		
240	pceadleenikgkkckirtpkiapvkfeligctsvktzakfcgvctdgrcctphrttt	299
Db		
241	PCEADLEENIKGKKCIRTPTKISKIPKEFLSGCTSMKT YRAKFCGVCTDGRCTPHRTT	300
QY		
300	lpvefcpdgeimknmmfiktcachyncpgdnndifeslyyrmygdma	348
Db		
301	LPVEFKPDGEVNKNMMFIKTCACHYNCPGDNDIFESLYRRMYGDMA	349
QY		

RESULT 4

R31599;	standard; Protein; 351 AA.
RC	R31599;
AC	24-MAR-1993 (first entry)
DE	Chicken nov protein.
OE	avian nephroblastoma; avian myeloblastoma virus; IGF binding site;
XW	insulin-like growth factor; Wilms' tumour.

RESULT 4

AC	R31599;	standard; Protein; 351 AA.
AC	R31599;	
AC	24-MAY-1993	(first entry)
DE	Chicken nov protein.	
DE	avian nephroblastoma; avian myeloblastoma virus; IGF binding site;	
KW	insulin-like growth factor; Wilm's tumour.	
OS	Gallus domesticus.	
OS		
FT	Key	Location/Qualifiers
FT	Peptide	1..24
FT	/label= signal peptide	
FT	/note= "only hydrophobic region of protein"	
FT	Binding_site	S6..63
FT	/label= IGF-binding site motif	
FT	/note= "corresponds to GCGCCXC consensus"	
FT	WO9300430-A.	
FT	07-JAN-1993.	
FT	25-JUN-1992; F00589.	
FT	25-JUN-1991; FR-007807.	
FT	(CNRS) CENT NAT RECH SCI.	
FT	Martinerie C, Perbal B;	
FT	WPI; 93-036377/04.	
FT	N-PSDB; Q36031.	
FT	Nucleotide sequences hybridising to regions of chicken nov gene -	
FT	useful as probes for detecting complementary sequences to	
FT	evaluate development and/or differentiation of tumours	
FT	Claim 1; Fig 1; 67pp; French.	
CC	This amino acid sequence was deduced from the nucleotide sequence	
CC	of a chicken nov gene clone isolated from a gene bank prepared from	
CC	chicken embryonic fibroblasts screened with a tumour-derived probe	
CC	The only hydrophobic region occurs within the putative signal	
CC	peptide suggesting that the protein is secreted. The protein also	
CC	contains the consensus motif of proteins which bind to insulin-like	
CC	growth factor. It is known that the human IGFII gene is	
CC	overexpressed in some Wilm's tumours and a similar deregulation of	
CC	IGFII expression could be involved in nephroblastoma development.	
CC	The deduced nov protein sequence contains 39 (non-clustered)	
CC	cysteine residues.	
CC	Sequence	351 AA.

Query Match 53.8%; Score 1460; DB 6; Length 351;
Best Local Similarity 54.2%; Pred. No. 4.62e-129;
Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;
b 3 tggggglpvlilllllllprcsvgreaacprcgcgcgaep-prcapvnpavldgcaccl 61

QY 2 TAASMG-PVRVAVFVLLALCSRPAVQNGSGPC--RCDEPAPRCPAGVSLVLDGCGCR 58
 Db 62 vcarqgescepllpdesggyldrdpdegagagcmvlegdncvfdgmylnqetfqp 121
 QY 59 VCAKQGLGELCTERDPCDPHKGLFCDFGSPANRKGIVCTAKDAPCIFGTVYRSGESFQS 118
 Db 122 skyqctdtdgqigclprnlglllpdpdpfprkievgececekwcd-prdevllggf 180
 QY 119 SKYQCTCLDGAAGVCMPLCSMDVRLSPDCPFPRRVKLPKGCCEWVCDPEKQTVVGP- 177
 Db 181 amaayqeatlgidvedsaancieqtewsacskcmgfsfrvtrnnrqcmvktqlc 240
 QY 178 ALAAYRLEDTFGDPFMIIRANCLVQTEWSACSKTCGMSITRVNDNASRLKQSRLC 237
 Db 241 mmrpeene-epedkkgkciqtckmkavrfeykntavqtkpkygclndgrcctphn 299
 QY 238 MVRPCEADLEENIKKGGKICIRPKIKFELSGCTSMKTYRAKFCGVCIDGRCCTPHR 297
 Db 300 tktiqvfcpggkflkpmmlintevchgnopqennaaffqpl 342
 QY 298 TTTLPVEFKPCDGEVKKNNMFIKTCACHYNCP-GDNDIFESL 339

RESULT 5
 ID R25565 standard; Protein; 379 AA.
 AC R25565;
 DT 18-JAN-1993 (first entry)
 DE Beta-IG-M1.
 KW Transforming growth factor beta; induced; CEF-10; v-src; chicken;
 KW embryo; fibroblasts; TGF-beta.
 OS Mus musculus.
 PN EP-495674-A.
 PD 22-JUL-1992.
 PF 17-JAN-1992; 300429.
 PR 18-JAN-1991; US-642291.
 PR 10-JAN-1992; US-816270.
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
 DR WPI; 92-243508/30.
 DR N-PSDB; Q26421.
 PT TGF-beta induced gene family - encodes proteins involved in
 PT growth and differentiation effects of TGF-beta-1
 PS Claim 2; Fig 1; 35pp; English.
 CC The protein sequence was deduced from the DNA sequence obtd. by
 CC screening a cDNA library made from AKR-2B mouse cells induced with
 CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
 CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
 CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
 CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
 CC Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
 CC induced by v-src in chicken embryo fibroblasts and is identical
 CC to the protein encoded by cyf61, an immediate early response gene
 CC induced in quiescent BALB 3T3 cells by serum treatment. Residues
 CC 49-56 of beta-IG-M1 conform to the GCGCXXC motif reported in the
 CC amino half of insulin-like growth factor (IGF) binding proteins.
 CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
 CC an amino acid sequence with strong homology to a motif found near the
 CC C-terminal of the malarial circumsporozoite (CS) protein, which is
 CC highly conserved among all species of malarial parasites sequenced

CC to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate
 CC cell-cell and cell-extracellular matrix interactions, such as
 CC properdin, thrombospondin, and tRAB. The proteins encoded by
 CC TGF-beta induced genes are likely to be involved in mediation of
 CC the biological effects of TGF-beta relating to cell growth and
 CC differentiation. See also R25566.
 SQ Sequence 379 AA;

Query Match 41.0%; Score 1113; DB 5; Length 379;
 Best Local Similarity 48.3%; Pred. No. 5.50e-95;
 Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;

Db 1 messtfrtlava-vtllhl-trials-tcpaachcple-apkcapgvglvrdgcckvc 56
 QY 1 MTAASMGFVRVAVFVLLALCSRPAVQNGSGPCRCDEPAPRCPAGVSLVLDGCGCRVC 60
 Db 57 akqinedcsktqpcdhtkglecnfgaastalkgicraqesgrpceynsrlvngesfqpn 116
 QY 61 AKQGLGELCTERDPCDPHKGLFCDFGSPANRKGIVCTAK-DGAPCIFGTVYRSGESFQS 119
 Db 117 khqctcidgavgcipclpcqelplnlgcpnprlvkvsqcccewcdedsikdlddd 176
 QY 120 CKYQCTCLDGAAGVCMPLCSMDVRLSPDCPFPRRVKLPKGCCEWVCDPE-KD----QT 173
 Db 177 dlglidaseveltrnneliagkgsalkrlpvfgtprvlfnplhahgkclvqttswsq 236
 QY 174 -VVG--PA---LA-----AY-R---LE--DTFGPDF-TM---IRAN---CLVQTTEWSA 208
 Db 237 cskacgtgistrvndnpeclrvketricvpcgpyvsslkkgkckskkspvrf 296
 QY 209 CSKTCGMSITRVNDNASCRLEKQSRLECMVRCEADLEENIKKGGKICIRPKIKRIF 268
 Db 297 tyagcsavkkyrpkycgscvdgrcctplqtrtvkmrfcedgemfaknmvliqscnyn 356
 QY 269 ELSGCTSMKTYRAKFCGVCIDGRCCTPHRTTLPVEFKPCDGEVKKNNMFIKTCACHYN 328
 Db 357 cphpneaf-rly 368
 QY 329 CPDNDI-FESLY 340

RESULT 6
 ID R90919 standard; Protein; 375 AA.
 AC R90919;
 DT 23-JUN-1996 (first entry)
 DE Connective tissue growth factor-2.
 KW CTGF-2; connective tissue growth factor-2.
 KW cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
 KW insulin-like growth factor; fibroblast growth factor; Cry61.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal_peptide
 FT Protein 25..375
 FT /label= mature_protein
 FN W09601896-A.
 PD 25-JAN-1996.
 PF 12-JUL-1994; U07736.
 PR 12-JUL-1994; W0-U07736.

PA	(HDMA-) HUMAN GENOME SCI INC.
PI	Adams MD, Li H;
PR	WPI; 96-097626/10.
DR	N-PSDB; T12653.
DR	Connective tissue growth factor-2 and DNA encoding it - useful to
PT	enhance the repair of connective and support tissue, and to enhance
PT	wound healing
PS	Claim 1: Fig 1A-C; 46pp; English.
PS	Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
CC	(T12653) isolated from a human foetal lung cDNA library. The CTGF
CC	polypeptides are structurally and functionally related to a family
CC	of growth factors which include IGF (Insulin-like growth factor),
CC	PDGF (platelet-derived growth factor), and FGF (fibroblast growth
CC	factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
CC	to Cry61. Cry61 is a growth factor-inducible immediate early gene
CC	initially identified in serum-stimulated mouse fibroblasts. It encodes
CC	a member of an emerging family of secreted proteins which are also a
CC	group of cysteine-rich proteins. This group of GFs are important for
CC	normal growth, differentiation, morphogenesis of the cartilaginous
CC	skeleton of an embryo and cell growth.
CC	Sequence 375 AA;
SQ	Sequence

Query Match	36.9%;	Score 1002;	DB 16;	Length 375;
Best Local Similarity	47.2%;	Pred. No. 3,886-84;		
Matches	175;	Conservative	71;	Mismatches 84; Indels 41; Gaps 21;
<hr/>				
Db	6	vrelalvlllhl-trvgls-tcpdchcple-apkcapgvglvrdgcgckvcakqlne	62	
Qy	9	VR-VAFVV-LTALCSRAVGNQSGPCRDEPAPRCPAGVSLVLDDGCCRCVCAKQLGE	66	
<hr/>				
Db	63	dorktpocdhktglecnfgaaastalkgicragesgrpcynariyvqngsfqpncnhqct	122	
Qy	67	LCTERDPCDPHKGFLCDFGSPANRKIGCVCTAK-DGAPCIFGGTVYRSSEFSQSCKYQCT	125	
<hr/>				
Db	123	cjgwirgaciplcqelshlnlpnbpnrivkvrtggccsewcdedsikdpmedgdgllgk	182	
Qy	126	CIDGNAG-CMPLCSMDVRLSPDCPPRRRVKLPGKCCEEWCD-----PK-QDT-WVGP	177	
<hr/>				
Db	183	glgfdaaseveltrnneliavkgsgskripvfqmeprrilypnlpggkciqvttawagcaak	242	
Qy	178	AL---AA---Y-R--L-E----DT-----FGPD-P-TM---IRA-NCLIVQTTEWSACSK	211	
<hr/>				
Db	243	tcggiatrvtndnpceerlvketricevrpgpvysalkkgkcsaktksapevfrftya	302	
Qy	212	TCGWGISTRVTNNDNASCRLEKQSRLCMVWRPCEADLEENIKGKKCIPTPKISKPIKFELS	271	
<hr/>				
Db	303	qclavkkvyrpkycgscvdgrccetpqrltrrvkmrfpedgetfknymmqisakncyneph	362	
Qy	272	GCTSMKYTRAKFCGVCTDGCCCTPHRTTLTPVEFKCPDGEVMKKNMMFIKTCACHYNCPG	331	
<hr/>				
Db	363	anea-a-fpyr	372	
Qy	332	DNDFESLYRY	342	

RESULT 7
ID R31608 standard; Protein; 205 AA.
AC R31608;
DT 24-MAY-1993 (first entry)
DE Homologous to chicken nov gene ex

KW	avian myeloblastoma; avian myeloblastoma virus;
KW	stringent hybridisation; ss.
FN	W09300430-A.
PD	07-JAN-1993.
PF	25-JUN-1992; F00589.
PR	25-JUN-1991; FR-007807.
FA	(CNRS) CENT NAT RECH SCI.
PI	Martinerie C, Perbal B;
DR	WPI; 93-036377/04.
PT	Nucleotide sequences hybridising to regions of chicken nov gene -
PT	useful as probes for detecting complementary sequences to
PT	evaluate development and/or differentiation of tumours
PS	Claim 21; Page 39; 67pp; French.
CC	The chicken nov gene is stimulated in avian nephroblastoma induced
CC	by avian myeloblastoma virus but not in normal adult kidney. A
CC	1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC	from chicken embryonic fibroblasts screened with a tumour-derived
CC	probe. Fragment XXII (Q36044) is part of the 3rd and 4th exons of
CC	the nov gene; nucleotide sequences which hybridise to Fragment XXII
CC	under stringent conditions (i.e. 50% formamide, 5 x SSC) are claimed.
CC	The claimed sequences preferably encode a protein with the sequence
CC	XXIII (R31608) or an amino acid sequence 60% homologous to it.
SQ	Sequence 205 AA;

Query Match	23.7%;	Score 644;	DB 6;	Length 205;
Best Local Similarity	98.8%;	Pred. No. 1.93e-49;		
Matches	81;	Conservative 1;	Mismatches 0;	Indels 0; Gaps 0;
Db	33	akdgapcifgtvtvrsgefsqscckycqctclddgavgcmlpsamdvlrpsdpcpfprvkl	92	
Qy	97	AKDGA PCIFG GTVTRSGESFQSSCKYQCTCLD GAVGCM LPSAMD VLRPSDCPFPRVKL	156	
Db	93	pgkccceewcdpekdqdtvlgpa	114	
Qy	157	PKGCCCEEWCDPEKDPQTWVGPA	178	

RESULT

0	Resou	ID	R31610 standard; Protein; 72 AA.
1		AC	R31610;
2		DT	24-MAY-1993 (first entry)
3		DE	Fragment XXVI with homology to chicken nov exon 3-encoded protein.
4		KW	avian myeloblastoma; avian myeloblastoma virus;
5		OS	stringent hybridisation.
6		OS	Homo sapiens.
7		PN	W09300430-A.
8		PD	07-JAN-1993.
9		PF	25-JUN-1992; F00589.
10		PR	25-JUN-1991; FR-007807.
11		PA	(CNBS) CENT NAT RECH SCI.
12		PI	Martinerie C, Ferbal B;
13		WP1	93-036377/04.
14		PT	Nucleotide sequences hybridising to regions of chicken nov gene -
15		PT	useful as probes for detecting complementary sequences to
16		PT	evaluate development and/or differentiation of tumours
17		PT	Claim 25; Page 40; 67pp; French.
18		CC	The chicken nov gene is stimulated in avian nephroblastoma induced
19		CC	by avian myeloblastoma virus but not in normal adult kidney. A
20		CC	1975bp cDNA sequence (036031) was isolated from a gene bank prepared
21		CC	from chicken embryonic fibroblasts screened with a tumour-derived

CC probe. Nucleotide sequences of the invention include those which
 CC code for a protein having at least 60% homology with the protein
 CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
 CC gene. In particular, the sequences of the invention encode a
 CC protein having the sequence of Fragment XXVI (R31610). (Sequence
 CC XXVI is described as a nucleotide sequence in the claims but is
 CC correctly described as an amino acid sequence in the disclosure).
 SQ Sequence 72 AA;

Query Match 21.3%; Score 579; DB 6; Length 72;
 Best Local Similarity 97.2%; Pred. No. 3.25e-43;
 Matches 70; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 dgapcifggtvyrsgesfqscqyqctcldgavgcmlcsmdvrlpsdpcpfrvklpg 60
 Qy 99 DGAPCIFGTVYRSGESFQSSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCPFRVVKLP 158
 Db 61 mceewwcvdepr 72
 Qy 159 KCCEEWVCDEPK 170

RESULT 9
 ID R31602 standard; Protein; 84 AA.
 AC R31602;
 DT 24-MAY-1993 (first entry)
 DE Chicken nov protein fragment IX.
 KW avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation.
 OS Gallus domesticus.
 PN W09300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C. Perbal B;
 DR WPI; 93-036377/04.
 DR N-PSDB; Q36035.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 8; Page 31; 67pp; French.
 CC The chicken nov gene is stimulated in avian nephroblastoma induced
 CC by avian myeloblastoma virus but not in normal adult kidney. A
 CC 1975bp cDNA sequence was isolated from a gene bank prepared from
 CC chicken embryonic fibroblasts screened with a tumour-derived probe.
 CC Fragment VIII (Q36035) is derived from the 3rd. exon of the nov gene;
 CC nucleotide sequences which hybridise to Fragment VIII under stringent
 CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
 CC sequences preferably encode a protein with at least 70% homology to
 CC amino acid sequence IX (R31602) which is encoded by Fragment VIII.
 SQ Sequence 84 AA;

Query Match 15.8%; Score 430; DB 6; Length 84;
 Best Local Similarity 57.7%; Pred. No. 4.30e-29;
 Matches 45; Conservative 19; Mismatches 13; Indels 1; Gaps 1;

Db 3 egdncvfdgmlyrnetqfcpackyqctcldgavgcmlcsmdvrlpsdpcpfrvklpg 62
 Qy 99 DGAPCIFGTVYRSGESFQSSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCPFRVVKLP 158

Db 63 eccekwvcd-prdevllg 79
 Qy 159 KCCEEWVCDEPKDQIVVG 176

RESULT 10
 ID R31609 standard; Protein; 71 AA.
 AC R31609;
 DT 24-MAY-1993 (first entry)
 DE Encoded by chicken nov gene exon 3 fragment.
 KW avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation.
 OS Gallus domesticus.
 PN W09300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C. Perbal B;
 DR WPI; 93-036377/04.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 25; Page 40; 67pp; French.
 CC The chicken nov gene is stimulated in avian nephroblastoma induced
 CC by avian myeloblastoma virus but not in normal adult kidney. A
 CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
 CC from chicken embryonic fibroblasts screened with a tumour-derived
 CC probe. Nucleotide sequences of the invention include those which
 CC code for a protein having at least 60% homology with the protein
 CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
 CC gene. (Sequence XXV is described as a nucleotide sequence in the
 CC claims but is correctly described as an amino acid sequence in the
 SQ Sequence 71 AA;

Query Match 15.2%; Score 412; DB 6; Length 71;
 Best Local Similarity 59.7%; Pred. No. 2.09e-27;
 Matches 43; Conservative 16; Mismatches 12; Indels 1; Gaps 1;

Db 1 egdncvfdgmlyrnetqfcpackyqctcldgavgcmlcsmdvrlpsdpcpfrvklpg 60
 Qy 99 OGAPCIFGTVYRSGESFQSSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCPFRVVKLP 158
 Db 61 eccekwvcd-pr 71
 Qy 159 KCCEEWVCDEPK 170

RESULT 11
 ID R31603 standard; Protein; 84 AA.
 AC R31603;
 DT 24-MAY-1993 (first entry)
 DE Polypeptide X homologous to chicken nov protein fragment.
 KW avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation.
 OS Homo sapiens.
 PN W09300430-A.
 PD 07-JAN-1993.

ID R31604 standard; Protein; 70 AA.
AC R31604; 1993 (first entry)
DE Chicken nov protein fragment XIV.
KW avian myeloblastoma; avian myeloblastoma virus;
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36038.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 13; Page 34; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences which hybridise to the 4th. exon of the nov gene;
CC Fragment XIII (Q36038) is derived from the 4th. exon of the nov gene;
CC nucleotide sequences which hybridise to Fragment XIII under stringent
CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC sequence preferably encode a protein with at least 86% homology to
CC amino acid sequence XIV (R31604) which is encoded by Fragment XIII.
SQ Sequence 70 AA;
Query Match 11.5%; Score 312; DB 6; Length 70;
Best Local Similarity 64.8%; Pred. No. 3.79e-18;
Matches 42; Conservative 9; Mismatches 14; Indels 0; Gaps 0;
Db 1 avrqeatlgldvdasancietwscakcmgfgstvtvtnnqqcmvktqlommr 60
QY 181 AVRLEDTFGDFTMIRANCLVOTTEWSACSKTCGMGISTRTVNDNASCRLKQSRLCMVR 240
Db 61 pcene 65
QY 241 PCEAD 245
RESULT 15
ID R31601 standard; Protein; 75 AA.
AC R31601;
DE 24-MAY-1993 (first entry)
KW Chicken nov protein fragment V.
KW avian nephroblastoma; avian myeloblastoma virus;
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to

PT evaluate development and/or differentiation of tumours
PS Claim 5; Page 28; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences which hybridise to Q36031 or specified
CC sub-fragments of it, under stringent conditions (i.e. 50% formamide,
CC 5 x SCC), are claimed. The claimed sequences preferably encode a
CC protein with amino acid sequence V (R31601).
SQ Sequence 75 AA;
Query Match 10.2%; Score 278; DB 6; Length 75;
Best Local Similarity 48.0%; Pred. No. 4.66e-15;
Matches 36; Conservative 14; Mismatches 22; Indels 3; Gaps 2;
Db 2 aatgrcpqpqgpcpatp-ptcagpgravidgcacclvcargqscsdlepdeasgly 60
QY 24 AVSQNCSGFC--RCPDEPAPRCAGVSLVLDGCGCCRCVCAKQJGELCTERDCD 81
Db 61 cdreadpanqtgict 75
QY 82 CDFGSPANKKIGVCT 96
Search completed: Wed Sep 17 09:31:56 1997
Job time : 91 secs.

maryh@stic

4383-1

NeWSprinter20

Wed Sep 17 11:21:24 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

 M P S R E L H
 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by IntelliGenetics, Inc.

MPsarch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:28:51 1997; MasPar time 13.03 Seconds
 773.831 Million cell updates/sec

Tabular output not generated.

Title: >OS-08-386-680-2
 Description: (1-349) from 5585270.pep
 Perfect Score: 2713
 Sequence: 1 MTAASMGFVRVAFVLLALC.....PGDNDFESLYRKMVGDMA 349

Scoring table: PAM 150
 Gap 11

Searched: 91006 seqs, 2888923 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pir51
 1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
 8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
 13:unann9 14:unann10 15:unann16:unrev

Statistics: Mean 42.720; Variance 82.202; scale 0.520

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2713	100.0	349	13	A40551	connective tissue gr 0.00e+00
2	2521	92.9	348	14	A40578	beta IG-M2 protein p 0.00e+00
3	2514	92.7	348	14	A53228	fisp-12 protein prec 0.00e+00
4	1460	53.8	351	13	S20078	NOV protein - chicke 3.42e-284
5	1415	52.2	357	13	I38069	gene novH protein - 2.53e-274
6	1152	42.5	375	13	A41428	CEF-10 protein precu 7.43e-217
7	1113	41.0	379	14	A35669	gene CxR61 protein p 2.27e-208
8	133	4.9	254	14	JCI1464	insulin-like growth 1.32e-06

9	133	4.9	254	14	I48599	insulin-like growth 1.32e-06
10	128	4.7	271	13	JC4584	insulin-like growth 8.02e-06
11	125	4.6	271	14	JCI463	insulin-like growth 2.33e-05
12	126	4.6	835	13	JP0076	nei protein - chicke 1.63e-05
13	123	4.5	238	13	B37252	insulin-like growth 4.71e-05
14	121	4.5	455	6	A43918	TGF-beta-related pro 9.49e-05
15	116	4.3	254	14	I48603	insulin-like growth 5.32e-04
16	117	4.3	277	13	I52825	gene MAC25 protein - 3.78e-04
17	117	4.3	282	13	S50031	prostacyclin-stimula 3.78e-04
18	118	4.3	1348	12	A43917	probable epidermal c 2.68e-04
19	118	4.3	1348	12	S27812	probable epidermal c 2.68e-04
20	115	4.2	272	6	A36082	insulin-like growth 7.48e-04
21	113	4.2	2813	3	VWU	von Willebrand facto 1.47e-03
22	110	4.1	263	6	S23009	insulin-like growth 4.01e-03
23	111	4.1	271	14	I48604	insulin-like growth 2.88e-03
24	111	4.1	304	6	A33274	insulin-like growth 2.88e-03
25	111	4.1	305	14	I48601	insulin-like growth 2.88e-03
26	111	4.1	305	6	JN0508	insulin-like growth 2.88e-03
27	110	4.1	3020	13	A43932	mucin 2 precursor, i 4.01e-03
28	109	4.0	124	6	S25113	insulin-like growth 5.59e-03
29	109	4.0	272	14	I48600	insulin-like growth 5.59e-03
30	106	3.9	258	14	A45403	insulin-like growth 1.49e-02
31	105	3.9	272	13	A53748	insulin-like growth 2.07e-02
32	106	3.9	807	14	A38152	F-spondin - rat 1.49e-02
33	102	3.8	259	2	IOH1	insulin-like growth 5.40e-02
34	102	3.8	291	2	JN0064	insulin-like growth 5.40e-02
35	104	3.8	310	6	A60967	insulin-like growth 5.40e-02
36	103	3.8	328	2	A41927	insulin-like growth 2.85e-02
37	103	3.8	500	4	F1WL6	I1 protein - human p 3.93e-02
38	102	3.8	798	7	A40526	integrin beta-7 chai 5.40e-02
39	104	3.8	3133	16	S70920	lectin - silkworm 2.85e-02
40	104	3.8	3133	16	S52093	lectin - silkworm 2.85e-02
41	101	3.7	388	7	A39756	circumsporozoite pro 7.42e-02
42	101	3.7	402	6	A45056	prepro osteogenic pr 7.42e-02
43	101	3.7	405	7	S05428	circumsporozoite pro 7.42e-02
44	101	3.7	412	3	OZ2QAF	circumsporozoite pro 7.42e-02
45	101	3.7	424	12	A54533	circumsporozoite pro 7.42e-02

ALIGNMENTS

RESULT	1
ENTRY	A40551
TITLE	connective tissue growth factor - human
ORGANISM	#formal_name Homo sapiens #common_name man
DATE	17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 25-Aug-1995
ACCESSIONS	A40551; S44205
REFERENCE	A40551
#authors	Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
#journal	J. Cell Biol. (1991) 114:1285-1294
#title	Connective tissue growth factor: a cysteine-rich mitogen secreted by human vascular endothelial cells is related to the SRC-induced immediate early gene product CEF-10.
#cross-references	UID:91373462
#accession	A40551
#molecule_type	mRNA
#residues	1-349
REFERENCE	S44205
#authors	Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.;

Luescher, T.F.
#submission submitted to the EMBL Data Library, April 1994
#description Differential cloning and expression of human connective
tissue growth factor.
#accession S44205
#status Preliminary
#molecule_type mRNA
#residues 1-349 ##label OEM
##cross-references EMBL:X78947
SUMMARY #length 349 #molecular-weight 38069 #checksum 8930

Query Match 100.0%; Score 2713; DB 13; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasmgprvafvllalcarpavqncgpcrcpdepaprcpagvslvldgcccrcvc 60
QY 1 MTAASMGPRVAFVLLALCSRPAVQNGSGPCRCDEPAPRCPCPAGVSLVLDGCGCCRCVC 60
Db 61 akqlgelcterdpchphkglfcdgfgspanrkiqvctakdgapciifgggtvyrsgesfgssc 120
QY 61 AKQLGELCTERDPCDPCHKGLFCDGFGSPANRKIGVCTAKDGACIFGGTVYRSGESFGSSC 120
Db 121 kyqctldgavcmplcsmdvrlpsdcpfprvklpgkceewvcdpdkdgtvvgpala 180
QY 121 KYQCTLDGAVGCMPLCSMDVRLPSDCPFPRVVKLPKGCCEWVCDPDKDQTVVGPALA 180
Db 181 ayrledtfgpdpmtiranclvqtewsacsktcgmgiatrvndnasrlekqsrlnmvr 240
QY 181 AYRLDTFGDPDMTIRANCLVQTEWSACSKTCGMGIATRVNDNASRLEKQSRCLMVR 240
Db 241 pceadleenikgkckirtpkiskpikfslsgctsmkyrakfgvctdgrcctphrttt 300
QY 241 PCEADLEENIKGKCKIRTPKISKPIKFLSGCTSMKYRAKFGVCTDGRCCTPHRTTT 300
Db 301 lpvefkcpdgvnmkmmfiktacachncpgdndifesylyrkmygdma 349
QY 301 LPVEFKCPDGEVNMKMMFIKTCACHNCPGDNDIFESLYYRKMYGDMA 349

RESULT 2
ENTRY beta IG-M2 protein precursor - mouse
TITLE
ORGANISM
DATE 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 30-Sep-1993
ACCESSIONS A40578
REFERENCE A40578
#authors Brunner, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
#journal DNA Cell Biol. (1991) 10:293-300
#title Identification of a gene family regulated by transforming growth factor-beta.
#cross-references MUID:91229699
#accession A40578
#status preliminary
#molecule_type mRNA
#residues 1-348 ##label BRU
SUMMARY #length 348 #molecular-weight 37794 #checksum 9188
Query Match 92.9%; Score 2521; DB 14; Length 348;

Best Local Similarity 91.1%; Pred. No. 0.00e+00;
Matches 318; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

Db 1 mlaavagpisalv-llalctrpatgdcacqcaaaahpcpagvslvldgcccrcvc 59
QY 1 MTAASMGPRVAFVLLALCSRPAVQNGSGPCRCDEPAPRCPCPAGVSLVLDGCGCCRCVC 60
Db 60 akqlgelcterdpchphkglfcdgfgspanrkiqvctakdgapcvggsvyrsgesfgssc 119
QY 61 AKQLGELCTERDPCDPCHKGLFCDGFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFGSSC 120
Db 120 kyqctldgavcmplcsmdvrlpsdcpfprvklpgkceewvcdpdkdrtavgpala 179
QY 121 KYQCTLDGAVGCMPLCSMDVRLPSDCPFPRVVKLPKGCCEWVCDPDKDQTVVGPALA 180
Db 180 ayrledtfgpdpmtiranclvqtewsacsktcgmgiatrvndntfcrlekqsrlnmvr 239
QY 181 AYRLDTFGDPDMTIRANCLVQTEWSACSKTCGMGIATRVNDNASRLEKQSRCLMVR 240
Db 240 pceadleenikgkckirtpkiskpikfslsgctsmkyrakfgvctdgrcctphrttt 299
QY 241 PCEADLEENIKGKCKIRTPKISKPIKFLSGCTSMKYRAKFGVCTDGRCCTPHRTTT 300
Db 300 lpvefkcpdgvnmkmmfiktacachncpgdndifesylyrkmygdma 348
QY 301 LPVEFKCPDGEVNMKMMFIKTCACHNCPGDNDIFESLYYRKMYGDMA 349

RESULT 3
ENTRY fisp-12 protein precursor - mouse
TITLE
ORGANISM fformal name Mus musculus #common name house mouse
DATE 19-May-1994 #sequence_revision 19-May-1994 #text_change 19-May-1994
ACCESSIONS A53228
REFERENCE A53228
#authors Ryseck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
#journal Cell Growth Differ. (1991) 2:225-233
#title Structure, mapping, and expression of fisp-12, a growth factor-inducible gene encoding a secreted cysteine-rich protein.
#accession A53228
#status preliminary
#molecule_type DNA
#residues 1-348 ##label RYS
##cross-references GB:M70641
GENETICS
#gene fisp-12
SUMMARY #length 348 #molecular-weight 37793 #checksum 9470
Query Match 92.7%; Score 2514; DB 14; Length 348;
Best Local Similarity 90.8%; Pred. No. 0.00e+00;
Matches 317; Conservative 21; Mismatches 10; Indels 1; Gaps 1;
Db 1 mlaavagpisalv-llalctrpatgdcacqcaaaahpcpagvslvldgcccrcvc 59
QY 1 MTAASMGPRVAFVLLALCSRPAVQNGSGPCRCDEPAPRCPCPAGVSLVLDGCGCCRCVC 60
Db 60 akqlgelcterdpchphkglfcdgfgspanrkiqvctakdgapcvggsvyrsgesfgssc 119
QY 61 AKQLGELCTERDPCDPCHKGLFCDGFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFGSSC 120

Qy 61 AKQGLGELTERDPOCPHKGFLCDFGSPANRKIGVCTAKGAPCIFGGTVYRSGESFQSSC 120
 Db 120 kyqctcidgavgevpplcsmdivrpspcdpfprvkvipgkckewcddepkdrtaavgpala 179
 Qy 121 KYQCTCLDGAAGCPLCSMDVRLSPDCPFPRVVKFLPGKCEEWCDDEPQDQTVVGPALA 180
 Db 180 ayrltdfsgdptmmranciqlvtwaeacaktgmigstrvntndntfcrlekqslcmvr 239
 Qy 181 AYRLTDFTGDPPTMIRANCLVQTTWASCKTGMIGSTRVNTNDNASCRLEKQSLCMVR 240
 Db 240 pceadleenikgkckirtpkiaipkvkfelsgtsvktyrakfcgvctdgroctphrttt 299
 Qy 241 PCEADLEENIKGKCKIRTPKISKP IKFELSGCTSMKTYRAKFCGVCTDGRCTPHRTTT 300
 Db 300 lpvefkcpdgeimkmmfiktctachynpcgndndifesiyyvkmvgdma 348
 Qy 301 LPVEFKCPDGEVKKMMF IKTCACHYNCEGNDNDFESIIYRKMYGDMA 349

RESULT 4
 ENTRY S20078 #type complete
 TITLE NOV protein - chicken
 ORGANISM #formal name Gallus gallus #common name chicken
 DATE 19-feb-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995

ACCESSIONS S20078
 REFERENCE S20078
 #authors Joliot, V.; Martinier, C.; Dambrine, G.; Plassiart, G.; Briac, M.; Crochet, J.; Perbal, B.
 #journal Mol. Cell. Biol. (1992) 12:10-21
 #title Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloblastosis-associated virus type 1-induced nephroblastomas.
 #cross-references MUID:92107157
 #accession S20078
 #status preliminary
 #molecule type mRNA
 #residues 1-351 #label JOL
 #cross-references EMBL:X59284

GENETICS
 #gene NOV
 #summary #length 351 #molecular-weight 38268 #checksum 4069

Query Match 53.8%; Score 1460; DB 13; Length 351;
 Best Local Similarity 54.2%; Pred. No. 3.42e-284;
 Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;

Db 3 tgggggplvllllllrpcevsagreaacprcggrcpaep-prcapgvpaavgcgcl 61
 Qy 2 TAASMG-PVRVAVVLLALCSRPVAVGNCSPGCRP-RCPEAPRCPAGVSLVLDGCGCCR 58
 Db 62 vcarqgescpplpdcseagglcdrgpedgggagimvlegdhcvcfdgmlyrnetfqp 121
 Qy 59 VCAKQGLGELTERDPCDHPKGLFCDFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQS 118
 Db 122 ackvgctordgiclpncnlgllpgdpdpfprkievpgceekwcd-prdevllggf 180
 Qy 119 SKYQCTCLDGAAGCPLCSMDVRLSPDCPFPRVVKFLPGKCEEWCDDEPQDQTVVGP- 177
 Db 181 amaaayrqeatigdvdsasancieqtetwasaacskcmgmfgatrvtrnngqcmvkvtrlc 240

Qy 178 ALAAAYRLDTFGDPPTMIRANCLVQTTWASCKTGMIGSTRVNTNDNASCRLEKQSLC 237
 Db 241 mmrpeene-epedkkgkckicqtkkamkavrfeykncvtavtykpryogclndgrcoctphn 299
 Qy 238 MYRPEADLEENIKGKCKIRTPKISKP IKFELSGCTSMKTYRAKFCGVCTDGRCTPHR 297
 Db 300 tktigvfcrcpgkflkppmmlintcvcgncpccpennaffqpl 342
 Qy 298 TTTLVPEFKCPDGEVKKMMF IKTCACHYNCP-GDNDIFESL 339

RESULT 5
 ENTRY I38069 #type complete
 TITLE gene novH protein - human
 ORGANISM #formal name Homo sapiens #common name man
 DATE 17-May-1996 #sequence_revision 17-May-1996 #text_change 17-May-1996

ACCESSIONS I38069
 REFERENCE I38069
 #authors Martinier, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, L.; Perbal, B.
 #journal Oncogene (1994) 9:2729-2732
 #title Structural analysis of the human nov proto-oncogene and expression in Wilms tumor.
 #cross-references MUID:94336229
 #accession I38069
 #status preliminary; translated from GB/EMBL/DBJ
 #molecule type DNA
 #residues 1-357 #label RES
 #cross-references EMBL:X78351; NID:g587422; CDS_PID:g825696

GENETICS
 #introns 28/3; 104/1; 188/1; 259/3
 #note gene name novH
 #summary #length 357 #molecular-weight 39162 #checksum 1640

Query Match 52.2%; Score 1415; DB 13; Length 357;
 Best Local Similarity 51.8%; Pred. No. 2.53e-274;
 Matches 177; Conservative 68; Mismatches 91; Indels 6; Gaps 5;

Db 18 ltflllhlqgvaatgrrcpqpcgrrcpatp-ptcapgvtravldgscclvcargaeac 75
 Qy 11 VAFVLLALCSRPVAVGNCSPGCRP-RCPEAPRCPAGVSLVLDGCGCCRCAKQGLGELC 68
 Db 76 adlepdcseagglcdreadpangtgcctavegdnvcvfdgviyrsgekfgpsckfgctord 135
 Qy 69 TERDPCDPHKGFLCDFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLD 125
 Db 136 ggqgcvrqcldvllpepncpaprkvvepgceekwgcgdeeds.lggltaayrpeatl 195
 Qy 129 GAVGCMPLCSMDVRLSPDCPFPRVVKFLPGKCEEWCDDEPQDQTVVGPALAAAYRLDTF 188
 Db 196 qvevsdasvncieqtetwasaacskcmgmfgatrvtrnngqcmvkvtrlc 255
 Qy 189 GPDTMIRANCLVQTTWASCKTGMIGSTRVNTNDNASCRLEKQSLCMVRPCEADLEE 248
 Db 256 ptdkkgkckirtkkslkaahlqfknctslhtykrfcgvcdgrcoctphntktiaefgc 315
 Qy 249 NI-KKGGKCKIRTPKISKP IKFELSGCTSMKTYRAKFCGVCTDGRCTPHRTTTLVPEFKC 307

[illegible]

TITLE	gene CYR61 protein precursor - mouse
ORGANISM	Formal name Mus musculus #common name house mouse
DATE	28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 18-Oct-1996
ACCSIONS	A35669; I48319
REFERENCE	A35669
#authors	O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.
#journal	Mol. Cell. Biol. (1990) 10:3569-3577
#title	Expression of cyr61, a growth factor-inducible immediate-early gene.
#cross-references	MUID:90287146
#accession	A35669
##status	preliminary
##molecule_type	mRNA
##residues	1-379 #label OAB
##cross-references	GB:M32490
##note	the authors translated the codon GAT for residue 379 Gln
REFERENCE	I48319
#authors	Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.
#journal	Nucleic Acids Res. (1991) 19:3261-3267
#title	Promoter function and structure of the growth factor-inducible immediate early gene cyr61.
#cross-references	MUID:91288203
#accession	I48319
##status	translated from GB/EMBL/DBJ
##molecule_type	DNA
##residues	1-379 #label RES
##cross-references	EMBL:X56790; NID:g50632; CDS_PID:g50633
GENETICS	
#gene	CYR61
#introns	21/3; 93/1; 208/1; 279/3
SUMMARY	#length 379 #molecular-weight 41709 #checksum 3726
Query Match	41.0%; Score 1113; DB 14; Length 379;
Best Local Similarity	48.3%; Pred. No. 2,276-208;
Matches	180; Conservative 79; Mismatches 76; Indels 38; Gaps :
Db	1 masstfrlava-vtlhl-trials-topaachople-apkcapgvglrvdgcgcvc 56
Qy	1 MTAASMGVVRVAVVLLALCSRPAGVQNGSGRCRCPDEPAPRCAPAGVSLVLDGGCCRCV 60
Db	57 akqlnedsktpcdhktgkfcnfgaastalkgicrdsgrpceynsrlyqngsqfn 116
Qy	61 AKQGLGELTERDPCDPHKGLCDFGSPANWKIGVCTAK-DGAPCIGGTVYRSGESFQSS 119
Db	117 khgctctidgavgcipicqalpalnlgcpmrlvkvaqgcceewcdedsakldadd 178
Qy	120 CKYQCTCLDGAVGCMPLCSMDVRLPSPDCFPFRVRVKLPKCKCEWVDE-P-KD----QT 173
Db	177 dlglgdaaeveltrmneliaigkgselkrlpvfgteprvlfnplhahgqkclvqtta 236
Qy	174 -AVG--FA--LA-----AY-R---LE---DTFGDPP-TM---IRAN---CLVQTTE 208
Db	237 cskscgtgistrvtdnnpceirlvketricevrcpgvpyysalkgkckaktkspepvrf 296
Qy	209 CSKTCGMGISTRTVNDNASCRLEKSGRLQWPCRADLEENIKGKKCIHTPKSPKPF 268
Db	297 tyagcsavkkyrpkycgscvdgrcctplqtrtvkmfrfcdgfmfknvmgscckcny 356

```

QY 269 ELSGCTSMKTYRAKFCVCGTDCRCCTPHRTTLTLPVEFKCDGEVKKMMFMFKTCACHYN 328

Db 357 cphpnasf-rly 368
|| | |
QY 329 CPGNDI-FESLY 340

RESULT 8
ENTRY JCL1464 #type complete
TITLE insulin-like growth factor-binding protein 4 precursor - rat
ORGANISM #formal name Rattus norvegicus #common name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
01-Dec-1995

ACCESSIONS JCL1464; A37252; E40403; A61120; B33570; A61118
REFERENCE
#authors Gao, L.; Ling, N.; Shimasaki, S.
#journal Biochem. Biophys. Res. Commun. (1993) 190:1053-1059
#title Structure of the insulin-like growth factor Binding protein-4
gene.
#accession JCL1464
##molecule_type DNA
##residues 1-254 ##label GAO
##cross-references GB:L08276

REFERENCE A37252
#authors Shimasaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.
#journal Mol. Endocrinol. (1990) 4:1451-1458
#title Molecular cloning of the cDNAs encoding a novel insulin-like
growth factor-binding protein from rat and human.
#cross-references MUID:91133415
#accession A37252
##molecule_type mRNA
##residues 1-254 ##label SHI

REFERENCE A40403
#authors Shimasaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
#journal J. Biol. Chem. (1991) 266:10646-10653
#title Identification of five different insulin-like growth factor
binding proteins (IGFBPs) from adult rat serum and
molecular cloning of a novel IGFBP-5 in rat and human.
#cross-references MUID:91244847
#accession E40403
##molecule_type protein
##residues 22-26,'X',28-29,'X',31-37,'X',39-40 ##label SH2

REFERENCE A61120
#authors Ceda, G.P.; Fielder, P.J.; Henzel, W.J.; Louie, A.; Donovan,
S.M.; Hoffman, A.R.; Rosenfeld, R.G.
#journal Endocrinology (1991) 128:2815-2824
#title Differential effects of insulin-like growth factor (IGF)-I
and IGF-II on the expression of IGF binding proteins
(IGFBPs) in a rat neuroblastoma cell line: isolation and
characterization of two forms of IGFBP-4.
#accession A61120
##molecule_type protein
##residues 'X',23-26,'X',28-29,'X',31-37,'X',39-43,'X',45,'X',47-48
##label CED
##note both glycosylated and nonglycosylated forms of this
protein were observed

REFERENCE A33570
#authors Shimmonaka, M.; Schroeder, R.; Shimasaki, S.; Ling, N.
#journal Biochem. Biophys. Res. Commun. (1989) 165:189-195
#title Identification of a novel binding protein for insulin-like

```

```

growth factors in adult rat serum.
#cross-references MWID:90073708
#accession B33570
##molecule type protein
##residues 22-26,'X',28-29,'X',31-37,'X',39-43,'X',45-52,'X',54,
'XX',57-58,'X',60-61 ##label SH3
REFERENCE
#A61118
#authors Cheung, P.T.; Smith, E.P.; Shimasaki, S.; Ling, N.;
Chernausek, S.D.
#journal Endocrinology (1991) 129:1006-1015
#title Characterization of an insulin-like growth factor binding
protein (IGFBP-4) produced by the B104 rat neuronal cell
line: chemical and biological properties and differential
synthesis by sublines.
#accession A61118
##molecule type protein
##residues 22-26,'X',28-29,'XX',32-33 ##label CHE
GENETICS
#117/1; 165/3; 210/3
#introns
#CLASSIFICATION #superfamily thyroglobulin type I repeat homology
glycoprotein
#FEATURE
1-21 #domain signal sequence #status experimental #label SIG
22-254 #product insulin-like growth factor-binding protein 4
#status experimental #label MAT
170-245 #domain thyroglobulin type I repeat homology #label
THY1
125 #binding site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY
#length 254 #molecular-weight 27745 #checksum 4589
Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No.1.32e-06;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
b 1 mlpfglaallagprpslqde-ahhpcpcseeklarcrrpvgceelrvpgcgccatc 59
| | : |||| :||| : : | :||| :||| : | :| : |||| |
2y 6 MGPVRVAFVLLALCSRAVQNGSGPCR-CPDEPAPRC-PA-GVS-LVLD-GCGGCRVC 60
| | : |||| :||| : : | :||| :||| : | :| : |||| |
b 60 aligmpc 67
| | | |
2y 61 AKQLGELC 68
| | | |
RESULT 9
ENTRY #type complete
TITLE insulin-like growth factor binding protein 4 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
18-Oct-1996
ACCESSIONS I48599; 538668
REFERENCE I48599
#authors Bethel, C.R.; Vitullo, J.C.; Miller, R.E.; Aron, D.C.
#journal Biochem. Mol. Biol. Int. (1994) 34:385-392
#title Molecular cloning of mouse insulin-like growth factor binding
protein 4 (IGFBP4) cDNA and expression of a fusion protein
with IGF-binding activity.
#cross-references MWID:95152444
#accession I48599
#status preliminary; translated from GB/EMBL/DDBV

```


##molecule_type DNA
##residues 1-254 ##label RES
##cross-references EMBL:X76066; NID:g416032; CDS_PID:g416033
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE 170-245
SUMMARY #domain thyroglobulin type I repeat homology #label THY1
#length 254 #molecular-weight 27807 #checksum 3509
Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No. 1.32e-06;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
Db 1 mlpfglvaallaaagrpaslgde-ahcpccseeklarcpvpgveelvrepqcccatc 59
Qy 6 MGPVRVAVVLLALCSRPVQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVC 60
Db 60 alglmpc 67
Qy 61 AKQLGELC 68

RESULT 10
ENTRY J04584 #type complete
TITLE Insulin-like growth factor binding protein-5 - pig
ORGANISM #formal name Sus scrofa domestica #common name domestic pig
DATE 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 03-Jun-1996
ACCESSION J04584
REFERENCE J04584
#authors White, M.E.; Diao, R.; Hathaway, M.R.; Mickelson, J.; Dayton, W.R.
#journal Biochem. Biophys. Res. Commun. (1996) 218:248-253
#title Molecular cloning and sequence analysis of the porcine insulin-like growth factor binding protein-5 complementary deoxyribonucleic acid.

#accession J04584
##molecule_type mRNA
##residues 1-271 ##label WHI
##cross-references GB:U41340
##experimental_source skeletal muscle
COMMENT This protein has essential roles in the regulation and coordination of insulin-like growth factors action. This protein enhances the in vitro activity of the insulin-like growth factors, plays a role during myoblast proliferation and differentiation, and is important in the growth and development of muscle tissue. differentiation; growth factor; skeletal muscle

KEYWORDS 1-19
FEATURE 20-271
SUMMARY #domain signal sequence #status predicted #label SIG
#product insulin-like growth factor binding protein-5
#status predicted #label MAT
#length 271 #molecular-weight 30323 #checksum 809

Query Match 4.7%; Score 128; DB 13; Length 271;
Best Local Similarity 37.5%; Pred. No. 8.02e-06;
Matches 30; Conservative 11; Mismatches 31; Indels 8; Gaps 7;

Db 2 vltavlllaacagpdlgsfvhcepcdekalsmcpplgclvkdpgcgccmtcala 61
Qy 9 VRVAVVLLALCSRPVQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVAKQ 63

Db 62 egscgqvtyerc-aq-glrc 79
Qy 64 LGELCTE-RDPCDPHKGFLFC 82

RESULT 11
ENTRY J01463 #type complete
TITLE Insulin-like growth factor-binding protein 5 precursor - rat
ORGANISM #formal name Rattus norvegicus #common name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995
ACCESSION J01463; A04003; F40403
REFERENCE J01463
#authors Zhu, X.; Ling, N.; Shimasaki, S.
#journal Biochem. Biophys. Res. Commun. (1993) 190:1045-1052
#title Cloning of the rat insulin-like growth factor binding protein-5 gene and DNA sequence analysis of its promoter region.

#accession J01463
##molecule_type DNA
##residues 1-271 ##label ZHU
##cross-references GB:L08275
REFERENCE A04003
#authors Shimasaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
#journal J. Biol. Chem. (1991) 266:10646-10653
#title Identification of five different insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human.

#cross-references M01D:9124847
#accession A04003
##molecule_type mRNA
##residues 1-271 ##label SHI
##cross-references GB:M62781
#accession F40403
#status preliminary
#molecule_type protein
#residues 20-25,'X',27-28,'X',30-36,'X',38-43,'X',45-51,'X',53,'XX' ##label SH2

GENETICS 112/3; 188/3; 228/3
#introns #superfamily thyroglobulin type I repeat homology
CLASSIFICATION
FEATURE 1-19
20-271 #domain signal sequence #status predicted #label SIG
#product insulin-like growth factor binding protein 5
#status predicted #label MAT
#domain thyroglobulin type I repeat homology #label MAT
#length 271 #molecular-weight 30298 #checksum 1075

191-262
SUMMARY Query Match 4.6%; Score 125; DB 14; Length 271;
Best Local Similarity 36.0%; Pred. No. 2.33e-05;
Matches 27; Conservative 12; Mismatches 28; Indels 8; Gaps 7;

Db 7 llllaacavpadlgsvfhcepcdekalsmcpplgclvkepgcgccmtcalaegsc 66
Qy 14 VLLALCSRPVQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVAKQLGELC 68

Db 67 gytyerc-aq-glrc 79
Qy 69 TE-RDPCDPHKGFLFC 82

```

RESULT 12
ENTRY   #type complete
TITLE   nel protein - chicken
ORGANISM #formal name Gallus gallus #common name chicken
DATE    24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change
01-Dec-1995
ACCESSIONS JP0076
REFERENCE   #formal name Gallus gallus #common name chicken
#authors   Taniguchi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
#journal   Dev. Dyn. (1995) 203:212-222
#title     New gene, nel, encoding a Mr 93K protein with EGF-like
#accession A38963
#molecule_type mRNA
#residues 1-835 #label MAT
#cross-references DBJ:D45365
#experimental_source 9-day embryo
REFERENCE   JP0076
#authors   Matsubashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
#submission submitted to JPIID, January 1995
#description A new gene, nel, encoding a Mr 93K protein with EGF-like
#accession JP0076
#molecule_type mRNA
#residues 1-835 #label MA2
#cross-references DBJ:D45365
#experimental_source 9-day embryo
CLASSIFICATION #superfamily von Willebrand factor type C repeat homology
FEATURE
273-333 #domain von Willebrand factor type C repeat homology
#label VWC\
395-592 #region EGF-like repeats
SUMMARY #length 835 #molecular-weight 93411 #checksum 7565
Query Match 4.6%; Score 126; DB 13; Length 835;
Best Local Similarity 40.3%; Pred. No. 1.63e-05;
Matches 25; Conservative 10; Mismatches 20; Indels 7; Gaps 6;
Db 274 ctmkgmtrefeswtgdgk-ncctmngtvtqcealicsldc-p-pnealsy-vd--gkcc 327
Qy 103 CIFGTVYRSGFSFOSCKYQCTCIDGAVGCMPL-CSMDVRLPSPDCFFPRVKLPKCC 161
Db 328 ke 329
Qy 162 EE 163

RESULT 13
ENTRY   #type complete
TITLE   insulin-like growth factor-binding protein 4 precursor -
ALTERNATE_NAMES human
#authors   IGFBP-4; inhibitory insulin-like growth factor-binding
#journal   protein; insulin-like growth factor-binding protein PC3
#formal name Homo sapiens #common name man
#accession 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change

```

```

06-Sep-1996
ACCESSIONS B37252; B39842; A36549; A60712; A54650; A49801; A34419
REFERENCE   A37252
#authors   Shimazaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.
#journal   Mol. Endocrinol. (1990) 4:1451-1458
#title     Molecular cloning of the cDNAs encoding a novel insulin-like
#accession #cross-references MUID:91133415
#molecule_type mRNA
#residues 1-258 #label SHI
#status not compared with conceptual translation
REFERENCE   A39842
#authors   Kiefer, M.C.; Maslarsz, F.R.; Bauer, D.M.; Zapf, J.
#journal   J. Biol. Chem. (1991) 266:9043-9049
#title     Identification and molecular cloning of two new 30-kDa
#accession #cross-references MUID:91225006
#molecule_type mRNA
#residues 1-258 #label KIE
REFERENCE   A36349
#authors   Latour, D.; Mohan, S.; Linkhart, T.A.; Baylink, D.J.; Strong,
#journal   D.D.
#title     Mol. Endocrinol. (1990) 4:1806-1814
#accession #cross-references MUID:91186988
#molecule_type mRNA
#residues 1-50, 'A', 52-197, 'F', 199-258 #label LAT
REFERENCE   A60712
#authors   Perkel, V.S.; Mohan, S.; Baylink, D.J.; Linkhart, T.A.
#journal   J. Clin. Endocrinol. Metab. (1990) 71:533-535
#title     An inhibitory insulin-like growth factor binding protein
#accession #cross-references MUID:91186988
#molecule_type mRNA
#residues 1-50, 'A', 52-197, 'F', 199-258 #label LAT
REFERENCE   A54650
#authors   Mohan, S.; Baylink, D.J.
#journal   Growth Regul. (1991) 1:110-118
#title     Evidence that the inhibition of TE85 human bone cell
#accession #cross-references MUID:93091814
#molecule_type protein
#residues 22-29, 'X', 28-29, 'X', 31-35 #label PER
REFERENCE   A54650
#authors   Mohan, S.; Baylink, D.J.
#journal   Growth Regul. (1991) 1:110-118
#title     Evidence that the inhibition of TE85 human bone cell
#accession #cross-references MUID:93091814
#molecule_type protein
#residues 22-29, 'X', 28-29, 'X', 31-35 #label PER
REFERENCE   A49801
#authors   Culouacou, J.M.; Shoyab, M.
#journal   Cancer Res. (1991) 51:2813-2819
#title     Purification of a colon cancer cell growth inhibitor and its

```


Db 1 mlpcslvaallltagrpisqde-aihcppcseeklarcrppvgeelvrepgccsaatc 59
||| : |||::: : |::: ||| : ||| : ||| :
QY 6 MGPIRVAFVLLALCSRPAVGQNCSGPCR-CPDEPAPRC-PA-GYS-LVID-GCGCCRCVC 60

Db	60	a	l	g	l	g	64
Qy	61	A	K	Q	L	G	65

Search completed: Wed Sep 17 09:30:07 1997
Job time : 76 secs.

maryh@stic

4384-1

NeWSprinter20

Wed Sep 17 11:22:18 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPERCH_PP

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligent Genetics, Inc.

MPERCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:27:42 1997; MacPar time 9.18 Seconds
Tabular output not generated. 806.556 Million cell updates/sec

Title: >US-08-386-680-2
Description: (1-349) from 5985270.pcp
Perfect Score: 2713
Sequence: 1 MTAASMGFVRVAFVLLALC.....PGNDIFESLYRKMGDMA 349

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 44.133; Variance 66.349; scale 0.665

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	2713	100.0	349	CTGF HUMAN	0.00e+00
2	2514	92.7	348	CTGF MOUSE	0.00e+00
3	1478	54.5	343	NOV XENLA	0.00e+00
4	1460	53.8	351	NOV CHICK	0.00e+00
5	1450	53.4	353	NOV COTJA	0.00e+00
6	1415	52.2	357	NOV HUMAN	0.00e+00
7	1152	42.5	375	CE10 CHICK	6.10e-277
8	1113	41.0	379	CYR61 MOUSE	4.24e-266
9	133	4.9	254	IBP4_RAT	5.18e-09

10	128	4.7	867	9	SSPO_BOVIN	SCO-SPONDIN (FRAGMENT	4.95e-08
11	125	4.6	271	5	IBP5_RAT	INSULIN-LIKE GROWTH F	1.88e-07
12	123	4.5	258	5	IBP4_HUMAN	INSULIN-LIKE GROWTH F	4.54e-07
13	121	4.5	435	1	60A_DROME	60A PROTEIN PRECURSOR	1.09e-06
14	116	4.3	254	5	IBP4_MOUSE	INSULIN-LIKE GROWTH F	9.35e-06
15	113	4.2	272	5	IBP1_RAT	INSULIN-LIKE GROWTH F	3.32e-05
16	113	4.2	2813	10	VWF_HUMAN	VON WILLEBRAND FACTOR	3.32e-05
17	110	4.1	263	5	IBP1_BOVIN	INSULIN-LIKE GROWTH F	1.16e-04
18	111	4.1	271	5	IBP5_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
19	111	4.1	304	5	IBP2_RAT	INSULIN-LIKE GROWTH F	7.65e-05
20	111	4.1	305	5	IBP2_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
21	110	4.1	2274	6	MUC2_HUMAN	MUCIN 2 (INTESTINAL M	1.16e-04
22	109	4.0	272	5	IBP1_MOUSE	INSULIN-LIKE GROWTH F	1.75e-04
23	106	3.9	258	5	IBP4_BOVIN	INSULIN-LIKE GROWTH F	5.92e-04
24	105	3.9	272	5	IBP5_HUMAN	INSULIN-LIKE GROWTH F	8.86e-04
25	106	3.9	807	4	FSPO_RAT	F-SPONDIN PRECURSOR.	5.92e-04
26	102	3.8	259	5	IBP1_HUMAN	INSULIN-LIKE GROWTH F	2.92e-03
27	102	3.8	291	5	IBP3_BOVIN	INSULIN-LIKE GROWTH F	2.92e-03
28	104	3.8	311	5	IBP2_CHICK	INSULIN-LIKE GROWTH F	1.32e-03
29	103	3.8	328	5	IBP2_HUMAN	INSULIN-LIKE GROWTH F	1.97e-03
30	103	3.8	500	10	VLI_HPV6B	MAJOR CAPSID PROTEIN	1.97e-03
31	102	3.8	798	5	ITB7_HUMAN	INTEGRIN BETA-7 SUBUN	2.92e-03
32	104	3.8	3133	5	HMCT_BOVMO	HEMOCYTIN PRECURSOR (1.32e-03
33	101	3.7	388	3	CSF_FLARE	CIRCUMSPOROZOITE PROT	4.32e-03
34	101	3.7	399	1	EM8A_MOUSE	BONE MORPHOGENETIC PR	4.32e-03
35	101	3.7	402	1	EMP8_HUMAN	BONE MORPHOGENETIC PR	4.32e-03
36	101	3.7	405	2	CSP_PLAFO	CIRCUMSPOROZOITE PROT	4.32e-03
37	101	3.7	412	2	CSP_PLAFA	CIRCUMSPOROZOITE PROT	4.32e-03
38	101	3.7	424	3	CSP_PLAFT	CIRCUMSPOROZOITE PROT	4.32e-03
39	100	3.7	442	3	CSP_PLAFW	CIRCUMSPOROZOITE PROT	6.38e-03
40	101	3.7	551	11	YD3_CAEEL	HYPOTHETICAL ZINC MET	4.32e-03
41	99	3.6	111	5	IBP5_BOVIN	INSULIN-LIKE GROWTH F	9.40e-03
42	99	3.6	291	5	IBP3_MOUSE	INSULIN-LIKE GROWTH F	9.40e-03
43	99	3.6	292	5	IBP3_RAT	INSULIN-LIKE GROWTH F	9.40e-03
44	98	3.6	806	5	ITB7_MOUSE	INTEGRIN BETA-7 SUBUN	1.38e-02
45	96	3.5	803	4	FSPO_XENLA	F-SPONDIN PRECURSOR.	2.95e-02

ALIGNMENTS

RESULT	1	CTGF HUMAN	STANDARD;	PRT;	349 AA.
ID	P29279;				
AC	01-DEC-1992 (REL. 24, CREATED)				
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.				
CN	CTGF.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;				
RX	MEDLINE; 91373462.				
RA	BRADHAM D.M., IGARASHI A., POTTER R.L., GROTEENDORST G.R.;				
RL	J. CELL BIOL. 114:1285-1294 (1991).				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;				

RX MEDLINE; 93187114.
 RA IGARASHI A., BRADHAM D.M., OKOCHI H., GROTEENDORST G.R.;
 RL J. DERMATOL. 19:642-643(1992).
 [3]
 RN SEQUENCE FROM N.A.
 RA OSMAR B.S., WERNER A., YANG Z., GARNIER J.M., GENTZ R., LDESCHER T.F.;
 RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
 CC HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
 CC MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
 CC -1- SDBNIT: MONOMER.
 CC -1- ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; M29234; G180924; -;
 DR EMBL; M56201; G266235; -;
 DR EMBL; X78947; G474934; -;
 DR EMBL; X40531; A40531.
 DR PIR; S44205; S44205.
 DR MIM; 121009; -;
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING; SIGNAL; ALTERNATIVE SPLICING.
 FT SIGNAL 1 21
 FT CHAIN 22 349
 FT DOMAIN 101 167
 FT DOMAIN 256 330
 FT DISULFID 256 293
 FT DISULFID 273 307
 FT DISULFID 284 323
 FT DISULFID 287 325
 FT DISULFID 292 329
 FT CARBOHYD 28 28
 FT CARBOHYD 225 225
 FT VARSPIC 172 198
 FT SEQUENCE 349 AA; 38069 MW; C2IE9662 CRC32;
 Query Match 100.0%; Score 2713; DB 3; Length 349;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 mtaasmpvrvafvllalcsrpavqncsgpcrcpdpaprcpagsvlvdgcgcrcv 60
 QY 1 MTAASMPVRVAFVLLALCSRPVAVQNCSGPCRCDEPAPRCPAGVSLVLDGCGCRVC 60
 Db 61 akqlgelcterdpchkglfcdgspanrkigvctakdgapcifggtvyrsgesfgsc 120
 QY 61 AKQGLGELCTERDPCPHKGLFCDFGSPANRKIGVCTAKDGAFCIFGGTVYRSGESFGSSC 120
 Db 121 kvqctldgavcmplcsmdvrlpsdpdpfprvklpgkcccewvcdpdkdtvvgpala 180
 QY 121 KVQCTCLDGAVCMPLCSMDVRLPSDPCFFPRVKLPKCCCEWVCDPDKDTVVGPALA 180
 Db 181 avrledtfgdpdmiranclvqtewsacsktcgmgiatrvtndnasrleqkqrlcmvr 240
 QY 181 AVRLEDTFGDPDMIRANCLVQTIEWSACSKTCGMGISTRVTNDNASRLEKQRLCMVR 240
 Db 241 pceadleeinkgkckirtpkiskpikfelsgctamktyrakfgvctdgrcctphrttt 300

QY 241 PCEADLEENIKGKCKIRTPKISKPIKFSLGCTSMKTYRAKFCGCTDGRCTPHRTTT 300
 Db 301 lpvefkcpdgvemkmmfiktacachyncpgndndifesylyrkmygdma 349
 QY 301 LPVEFKCPDGEVEMKMMFMFKTACHYNCPGNDNDFESLYRYKMYGDMA 349
 RESDLT 2
 ID CTGF MOOSE STANDARD; PRT; 348 AA.
 AC P29268;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE CONNECTIVE TISSUE GROWTH FACTOR PRECDRSOR (CTGF) (FISP-12 PROTEIN).
 GN CTGF OR FISP12 OR FISP-12.
 OS MUS MUSCULUS (MOOSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91363290.
 RA RYSECK R.-P., MACDONALD-BRAVO H., MATTEI M.-G., BRAVO R.;
 RL CELL GROWTH DIFFER. 2:225-233(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91229699.
 RA BRUNNER A., CHINN J., NEDEBADER M.G., PURCHIO A.F.;
 RL DNA CELL BIOL. 10:293-300(1991).
 CC -1- INDUCTION: BY GROWTH FACTORS.
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; M70641; G193314; -;
 DR EMBL; M70642; G193316; -;
 DR EMBL; M80263; G201946; -;
 DR PIR; A53228; A53228.
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 21
 FT CHAIN 22 348
 FT DOMAIN 100 166
 FT DOMAIN 255 329
 FT DISULFID 255 292
 FT DISULFID 272 306
 FT DISULFID 283 322
 FT DISULFID 286 324
 FT DISULFID 291 328
 FT CONFLICT 161 161
 FT SEQUENCE 348 AA; 37793 MW; EAB92BE0 CRC32;
 Query Match 92.7%; Score 2514; DB 3; Length 348;
 Best Local Similarity 90.8%; Pred. No. 0.00e+00;
 Matches 317; Conservative 21; Mismatches 10; Indels 1; Gaps 1;
 Db 1 mlaavagplslav-llalctrpatgdcacqcaacaaaphcpagvslvdgcgcrcv 59

QY 1 MTAASMGPRVAVFVLLALCSRPAGVQNCSPCEPDEPAPCPAGVSLVLDGCGCCRV 60
 Db 60 akqigelcterdpdphkglfcdfgspanrkigvctakdgapcvfgsvvrsesqsc 119
 QY 61 AKQIGELCTERDPDHPKGLFCDFGSPANRKIGVCTAKDGAPCFIGGTVYRSGESFQSSC 120
 Db 120 kyqcteldgavgcplcsmdvrlpsdpcpfrvrkpgkckewwdekdrtavgpala 179
 QY 121 KYQCTELDGAVGCMPLCSMDVRLPSPDCPPFRVRKLPKGCCEEWVDEPKDQTVVGPALA 180
 Db 180 ayrltdtfgpdtmmranclvqtewasactcmgmlstrvndntforlekqsrclmvr 239
 QY 181 AYRLDTFGPDTMIRANCLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRCLMVR 240
 Db 240 pceadleenkkgkckirtkpkakpvkfelsctsvktvyrakfcgvctdgrctcphrttt 299
 QY 241 PCEADLEENIKKGKCKIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCTCPTHRTT 300
 Db 300 lpvefkpcdgeimkmmfiktacachynpcgndndifeslyrkmygdma 348
 QY 301 LPVEFKPCDGEVMMKMMFIKTCACHYNCPGNDNDFESLYRMYGDMA 349

RESULT 3
 ID NOV XENLA STANDARD; PRT; 343 AA.
 AC P31609;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NOV PROTEIN HOMOLOG PRECURSOR (XNOV).
 GN XNOV
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
 RN [1]
 RP YING Z., KING M.L.;
 RA SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -!- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; G37063; G1176415; -;
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 18
 FT CHAIN 19 343
 FT DOMAIN 93 159
 FT DOMAIN 249 323
 FT DISULFID 249 286
 FT DISULFID 266 300
 FT DISULFID 277 316
 FT DISULFID 280 318
 FT DISULFID 285 322
 FT CARBOHYD 285 265
 SQ SEQUENCE 343 AA; 38070 MW; BDF3BCA4 CRC32;
 Query Match 54.5%; Score 1478; DB 6; Length 343;
 Beat Local Similarity 53.7%; Pred. No. 0.00e+00;
 Matches 180; Conservative 69; Mismatches 78; Indels 8; Gaps 8;

Db 1 mtp-hlalcfill-1-lqqvasckcpeqcdqcpcep-pecapavslilidgcqccpvcarg 56
 QY 6 MGEVRVAVFVLLALCSRPAGVQNCSPCEPDEPAPCPAGVSLVLDGCGCCRVAKQL 64
 Db 57 gescshlnpcqcdkglcyefnadmetgtmalnegnacvfdgvvynreafqpsckyh 116
 QY 65 GELCTERDPDHPKGLFCDFGSPANRKIGVCTAKDGAPCFIGGTVYRSGESFQSSCKYQC 124
 Db 117 tclnhgicvpcrnldlllpgdpcpfrvrkvpgecekwcd-skeemaigfgfanaayr 175
 QY 125 TCLDGAVGCMPLCSMDVRLPSPDCPPFRVRKLPKGCCEEWVDEPKDQTVVGP-ALAAAYR 183
 Db 176 peatlgidasstsfaciagttewasactcmgmsvtrnrcarnemqkqiricmvrsee 235
 QY 184 LEDTFGPDTMIRANCLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRCLMVRPCE 243
 Db 236 eepgwhvekkgkcvrvrkttkpihfyknctsvqpkpfcgqcdgrctcphaktmh 299
 QY 244 ADLEENI-KKGKCKIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCTCPTHRTTIP 302
 Db 296 vefvcpdkrvkvpvmvstcvcchynpcqdesllq 330
 QY 303 VEFKPCDGEVMMKMMFIKTCACHYNCPGNDNDFE 337

RESULT 4
 ID NOV CHICK STANDARD; PRT; 351 AA.
 AC P28686;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NOV PROTEIN PRECURSOR.
 GN NOV.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BROWN LEIGHORN;
 RX MEDLINE; 92107157.
 RA MALOISEL V., MARTINERIE C., DAMERINE G., PLASSIART G., BRISAC M.,
 RA CROCHET J., PERBAL B.;
 RL MOL. CELL. BIOL. 12:10-21(1992).
 CC -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
 CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION
 CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -!- DEVELOPMENTAL STAGE: MAVI-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 CC ADULT KIDNEY.
 CC -!- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
 CC MUSCLE AND INTESTINE, IN THE EMBRYO. LUNG AND LESS SO IN BRAIN AND
 CC SPLEEN, IN ADULT CHICKEN.
 CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -!- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; X59284; G63703; -;

DR	PIR; S20078; S20078.
DR	PROSITE; PS00222; IGF BINDING.
KW	PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
FT	SIGNAL 1 24 POTENTIAL.
FT	CHAIN 25 351 NOV PROTEIN.
FT	DOMAIN 104 170 VMFC.
FT	DOMAIN 258 332 CTCK.
FT	DISULFID 258 295 BY SIMILARITY.
FT	DISULFID 275 309 BY SIMILARITY.
FT	DISULFID 286 325 BY SIMILARITY.
FT	DISULFID 289 327 BY SIMILARITY.
FT	DISULFID 294 331 BY SIMILARITY.
FT	CARBOHYD 274 274 POTENTIAL.
SQ	SEQUENCE 351 AA; 38268 MW; C70404065 CRC32;

Query Match	53.8%;	Score 1460;	DB 6;	Length 351;
Best Local Similarity	54.2%;	Pred. No. 0.00e+00;		
Matches 186;	Conservative	69;	Mismatches 80;	Indels 8; Gaps 7;

Ddb		3	tggggglpvlllllllprcevsgraeacprcpggrcraep-prcapgvpauidvgccl	61
Oy		2	TAA\$MG-PVRVAFVLLIACSRAPVQNGSGFC--RCDEPAPRCPAGSVLIVDGGCCGR	58
Ddb		62	vcarqrescspllpodesggyldrqpdeggggagimvlsgdnvcfdgmivngtcfq	121
Oy		59	VCAQLGELTERDCPDHKGFLCFDFGSPANKIGVCTAKDGAPCIFGTGYRSGESFOS	118
Ddb		122	skvqcctcdggqgciprcnlgllppgcdpfkrievgpccckwcd-prdevillgf	180
Oy		119	SCKYQCTCIDGAVGCMPCLSMVDRLFPSPDCFFPRVKLPKGCCEWVDEDFKDQTVVGP-	177
Ddb		181	amaayrqeatigdvdsaaancleqtewscascakmgfatrvtnnnqqcmvktrlc	240
Oy		178	ALAAYLEDTFGDPDMIRANCLVQTTEWSACSKTCGMGTSTRVNDNASCRLEKQSRLC	237
Ddb		241	mnrpcene-epsdkdkkciqttkamkavrfeyknctasvtqykrycgldndgrectphn	299
Oy		238	MVRPEADLEENIKKKCIITPKISKPIKFLESCTSMTKYRAKTCGVTDGRCCTPHR	297
Ddb		300	tktiqvfcrcpqgkfikkpmmlintcvhncpdcnaaffql	342
Oy		298	TTTTLPVEFKDGEVMKKNMFIKTCAHYNCP-GDNDIFESL	339

RESULT	NOV 5	STANDARD;	PRT;	353 AA.
ID	NOV COTVA			
AC	F42642;			
DT	01-NOV-1995	(REL. 32, CREATED)		
DT	01-NOV-1995	(REL. 32, LAST SEQUENCE UPDATE)		
DT	01-OCT-1996	(REL. 34, LAST ANNOTATION UPDATE)		
DE	NOV PROTEIN PRECURSOR.			

COFURNIX COTURNIX JAPONICA (JAPANESE QUAIL).
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
GALLIFORMES.
(1)
SEQUENCE FROM N.A.
WEISKIRCHEN R., BISTER K.;
SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
-1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL

1-1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL

CC	GROWTH REGULATION (BY SIMILARITY).
CC	-!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
CC	TO THE CEF-10/CYR61/CTGF/FIS-12/NOV PROTEIN FAMILY.
CC	-!- SIMILARITY: CONTAINS A VWFC DOMAIN.
CC	-!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC	EMBL: U13063; G532697; -.
DR	PROSITE; PS00222; IGF_BINDING.
DR	PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
KW	POTENTIAL.
FT	SIGNAL 1 26
FT	CHAIN 27 353
FT	DOMAIN 106 172
FT	DOMAIN 260 334
FT	DOMAIN 260 297
FT	DISULFID 277 311
FT	DISULFID 288 327
FT	DISULFID 291 329
FT	DISULFID 296 333
FT	CARBOHYD 276 276
FT	POTENTIAL.
FT	SEQUENCE 353 AA; 38667 MW; C4F5928D CRC32;
CSQ	

Query Match 53.4%; Score 1450; DB 6; Length 353;
Best Local Similarity 54.7%; Pred. No. 0.00e+00;
Matches 185; Conservative 66; Mismatches 78; Indels 9; Gaps 8;

10 pvl|||||lllllrrpsevnreapcprrcggrcpcapc-prcapgvpavldgcgcclvcarq 68
8 pvravfvl|llalcsreavg-QN-CSGPC---RCPDEPAPRCPAGVSLVLDGCGCGRCAQ 63
69 rgeacaplpcdesgglvcdrrpedgggtgicwlegcdnvcfdgmlyrngetfpacpekyy 128
64 LGELCTERDPCDGHGLFCDFGSPANRRIGVCTAKDGAFCIGTGVYRSGESFQSSCKY 123
129 ctcrdggicaplrcnlllpdpcpcpforkievpgceceklwv--eprdevlvggfamaay 187
124 CTCLDGAVGCMPLCSMDVRLSPDPCFFPRRVKLFSGKCEEWVDEPKDQTVVGP--ALAAY 182
188 rgeatlgldvdsasancleqtetwaaacarscmgmfstvrnrrnqgcemkvqtrlcmrrpc 247
183 RLEDTGPDPTMIRANCLVQTTEWSACSCKTCMGISTRVTNDNASRLKFKSRLCWVRPC 242
248 ene-epsdkkgkckirtkksmkavfeykntcsvqtkryvcgicldngdrccctphntktig 306
243 EADLEENIKKKKKCIRTPKISKPIKFELSGCTSMKTYRAFCGVCTDGRCTPHRTTILP 302
307 vefrcpqgkflk|kmmllntvcvchgnpcqsnnaifqpl 344
303 VEPKCPDGEVMKKNMWF|TKTACHYNCP-GNDIFESL 339

RESULT	6	STANDARD;	PRT;	357 AA.
AD	NOV HUMAN			
IC	P48745;			
DDT	01-FEB-1996	(REL. 33, CREATED)		
DDT	01-FEB-1996	(REL. 33, LAST SEQUENCE UPDATE)		
DDT	01-OCT-1996	(REL. 34, LAST ANNOTATION UPDATE)		
DE	NOV PROTEIN	HOMOLOG PRECURSOR (NOVH).		

HOMO SAPIENS (HUMAN).
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
EUTHERIA; PRIMATES.

RN [1] SEQUENCE FROM N.A.
 RP TISSUE=PLACENTA;
 RX MEDLINE; 94336229.
 RA MARTINERIE C., HOFF V., JOUBERT I., BADZIDCH M., SAUNDERS G.,
 RL STRONG L., PERBAL B.;
 CC ONCOGENE 9:2729-2732(1994).
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE
 CC STROMAL TYPE.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; X78351; G825696; --
 DR EMBL; X78352; G825696; JOINED.
 DR EMBL; X78353; G825696; JOINED.
 DR EMBL; X78354; G825696; JOINED.
 DR EMBL; X96584; E228691; --
 DR MIM; 164958; --
 KW PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 27
 FT CHAIN 28 357
 FT DOMAIN 108 174
 FT DDMAIN 264 338
 FT DISULFID 264 301
 FT DISULFID 281 315
 FT DISULFID 292 331
 FT DISULFID 295 333
 FT DISULFID 300 337
 FT CARBDHYD 97 97
 FT CARBOHYD 280 280
 FT SEQUENCE 357 AA; 39162 MW; DA8B009D CRC32;
 Query Match 52.2%; Score 1415; DB 6; Length 357;
 Best Local Similarity 51.8%; Pred. No. 0.00e+00;
 Matches 177; Conservative 68; Mismatches 91; Indels 6; Gaps 5;
 Db 18 ltf1-llhllgvaatqrcppqgpcratp-ptcagvtravldgcaccivcarqgsc 75
 Qy 11 VAFVLLALCSRPAGVQNCSPC--RCPEAPRCPAGVSVLDGCGCRVCAKQGLGELC 68
 Db 76 sdlepcdesgldraadpntqgictavdgndvfdgviyrzgekfqsckfctcd 135
 Qy 69 TERDPCDPHKLFCDFGSPANRKTGVCCTAKDAPCFIFGTVYRSGESFQSCYQCTCLD 128
 Db 136 ggqgcvprcqldvllpepcpapkvevpgcecekwldcgpedealggiltlaayrpeatl 195
 Qy 129 GAVGCMPLCSMDVRLSPDCFFPRVRKLPKGCCEWVCDPEKOOTVVGPAALAYRLEDTF 188
 Db 196 gvevdsavncieqtetwtacskacgmfgstrvtnnrqemlkqtrclmvrpceqeq 255
 Qy 189 GPDEPTMIRANLVQTEWSACSKTCGMGISTRTVNDNASCRLEKQSLRCWVPCADLEE 248
 Db 256 ptdkkgkclrtkslkhllqfknctslhlykprfgvcsdgrcctphntkqieafgc 315
 Qy 249 NI-KKGGKKCIINTPKISKPIKFIKLSGCTSMKTYRAKFCGVCTDRCCTPHRTTTLPEFKC 307
 Db 316 abqgikvkvpmvtdtctchncpknneafiqelelkttrckm 357

Qy 308 PDGEVMMKNNMTIKTCACHYCPGNDIF-ESLYYRKMGM 348
 RESULT 7
 ID CE10 CHICK STANDARD; PRT; 375 AA.
 AC P19336;
 DT 01-NDV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE CEF-10 PROTEIN PRECURSOR.
 OS GALLUS GALLUS (CHICKEN).
 DC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89145206.
 RA SIMMONS D.L., LEVY D.B., YANNONI Y., ERIKSON R.L.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:1178-1182(1989).
 CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
 CC -1- INDUCTION: BY V-SRC.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DDMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; J04496; G211436; --
 DR PIR; A41428; A41428.
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 22
 FT CHAIN 23 375
 FT DOMAIN 98 164
 FT DDMAIN 281 355
 FT DISULFID 281 318
 FT DISULFID 298 332
 FT DISULFID 309 348
 FT DISULFID 312 350
 FT DISULFID 317 354
 FT SEQUENCE 375 AA; 40651 MW; 68B4BC92 CRC32;
 Query Match 42.5%; Score 1152; DB 2; Length 375;
 Best Local Similarity 48.5%; Pred. No. 6.10e-277;
 Matches 176; Conservative 76; Mismatches 77; Indels 34; Gaps 19;
 Db 7 rpalaal-lclarlalgaapcavcqp-aapcagvglyvpgcgckvcaqlnedc 64
 Qy 10 RVAFVLLALC-SRPAGVQNCSPCPCPEAPRCPAGVSVLDGCGCRVCAKQGLGELC 68
 Db 65 artqcdhtkglecnfgaspaatngicraqsegrpcynskiyngesfqpnkhtqctci 124
 Qy 69 TERDPCDPHKLFCDFGSPANRKTGVCCTAK-DGAPCFIFGTVYRSGESFQSCYQCTCL 127
 Db 125 dgavgcipcpqelslpnlqpcsprlvkvpqgceewvcdskdalelegffekfegld 184
 Qy 128 DGAVGCMPLCSMDVRLSPDCFFPRVRKLPKGCCEWVCDPEKD--QTVVG--P---AL- 179
 Db 185 aasegeltrneliaivkgglkmlpvfgepqsrafcenpkciqvttswsqscctqgtglt 244
 Qy 180 AA---Y-R---L-E-----TFGPDPTM-I-RANCLIVQTENSACSKTCGMGIST 219

Ob 245 rvtndpckliketicevrrpqqpayaalkkqkktkktkspvrrfytgcsavkky 304
 Qy 220 RVTNDNASCRLKSRVLRCPADLEENIKRKGKCIIRPKISKIRKFSLGCTSMKTY 279
 Db 305 rpykycgcvdgrcttqgttvtvkirfcdcdgetftkavmmiqecrcnynpcnphaneay-p- 362
 Qy 280 RAKFCGCTGCTGCTPHRTITLVEFKCPOGEVMMKNMFIKTACHYNCPCNONIFESL 339
 Db 363 fyr 365
 Qy 340 YIR 342

RESULT 8 STANOARD; PRT; 379 AA.
 AC P18406;
 DT 01-NOV-1990 (REL. 16, CREATEO)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPOATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPOATE)
 DE CYR61 PROTEIN PRECURSOR (3CH61).
 GN CYR61.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ROENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C / 3T3;
 RX MEDLINE; 90287146.
 RA O'BRIEN T.P., YANG G.P., SANOERS L., LAU L.F.;
 RL MOL. CELL. BIOL. 10:3569-3577(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AJ; TISSUE=EMBRYONIC FIBROBLAST;
 RX MEDLINE; 91288203.
 RA LATINKIC B.V., O'BRIEN T.P., LAU L.F.;
 RL NUCLEIC ACIDS RES. 19:3261-3267(1991).
 CC -|- FUNCTION: CYR61 MIGHT ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS.
 CC -|- OEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MIO-G(1) IN
 CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
 CC -|- INDUCTION: BY GROWTH FACTORS.
 CC -|- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
 CC ANO OVARY, MOODERATE IN HEART, UTERUS, ANO SKELETAL MUSCLE, HIGHEST
 CC IN LONG.
 CC -|- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS ANO
 CC TO THE CEP-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -|- SIMILARITY: CONTAINS A VWFC COMAIN.
 CC -|- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; M32490; G309206; -.
 DR EMBL; X56790; G50633; -.
 OR PIR; A35669; A35669.
 OR PROSITE; P50022; IGF BINDING.
 KW GROWTH FACTOR BINDING, SIGNAL.
 FT SIGNAL 1 24
 FT CHAIN 25 379
 FT CYR61 PROTEIN.
 FT DOMAIN 98 164
 FT DOMAIN 284 358
 FT OISULFIO 284 321
 FT OISULFIO 301 335
 FT DISULFID 312 351

FT OISULFIO 315 353 BY SIMILARITY.
 FT OISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MW; 116B80C7 CRC32;
 Query Match 41.0%; Score 1113; OB 3; Length 379;
 Best Local Similarity 48.3%; Pred. No. 4.24e-266;
 Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;

Ob 1 masstfrtlava-vtllhl-trllals-tcpaachople-apkcapgvglvrdgcgckvc 56
 Qy 1 MTAASMGVPRVAFVLLALCSRPAVGQNCSPGCRCPDEPAPRCRCPAGVSLVLOGGCRVC 60
 Ob 57 akqlnedcsktqcdhtkglecnfgasataikgicraqsegirpceyhsrlyvqgesfqp 116
 Qy 61 AKQLGELCTERDPCDPHKGFLFCFGSPANRKGIVCTAK-OGAPCIFGTVYRSGESFQSS 119
 Ob 117 ckhqctcidgavgcipclpcqelslpnlgcpnprlvkvgqcccewcdedsikdsiddq 175
 Qy 120 KRYQCTCLOGAVGCMPLCSMDVRLPSPQPPFRVRLKLPGRCEEWCCE-P-KD-----QT 1
 Ob 177 dlglldaseveltrnnelialigkgsikrlpvgfctprvlnfnphagkqciqvgttaweq 236
 Qy 174 -VVG--PA---LA-----AY-R---LE--OTFGDP-TM---IRAN---CLVQTTEWSA 208
 Db 237 cskscgtgistrvtndnpeclrvketricevrrpqqpvyaslkkgkckskkspvrrf 296
 Qy 209 CSKTGCMGISTRVNONASCRLEKQSLRQWVRCEADLEENIKRKGKCIIRPKISKIRK 268
 Ob 297 tyagssvkvrykycgscvdrcttqltrtvkmrfrcecdgemfknmmigscckcny 356
 Qy 269 ELSGCTSMKTYRAKFCGCTGCTPHRTITLVEFKCPOGEVMMKNMFIKTACHY 328
 Ob 357 cphneaf-rly 368
 Qy 329 CFGONOI-FESLY 340

RESULT 9
 TO IBP4 RAT STANDARD; PRT; 254 AA.
 AC P21744;
 OT 01-MAY-1991 (REL. 18, CREATEO)
 OT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPOATE)
 OT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPOATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
 OE (IBP-4) (IGF-BINDING PROTEIN 4).
 GN IGFBP4 OR IGFBP-4.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ROENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEOLINE; 91133415.
 RA SHIMASAKI S., UCHIYAMA F., SHIMONAKA M., LING N.;
 RL MOL. ENOCROINOL. 4:1451-1458(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-OWALEY;
 RX MEOLINE; 93176147.
 RA GAO L., LING N., SHIMASAKI S.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:1053-1059(1993).

[3]	SEQUENCE OF 22-61.	
RRP	TISSUE-SERUM;	
RC	MEDLINE; 90073708.	
RR	SHIMONAKA M., SHIMODA R., SHIMASAKI S., LING N.;	
RR	BIOCHEM. BIOPHYS. RES. COMMON. 163:189-195(1989).	
RL	-1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs	
CC	AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH	
CC	PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE	
CC	INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.	
CC	-1- BINDS IGF-II MORE THAN IGF-I.	
CC	-1- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.	
CC	EMBL; L08276; E73337; --	
CC	PIR; B33570; B33570.	
DR	PIR; A37252; A37252.	
DR	PIR; JC1464; JC1464.	
DR	PIR; E40403; E40403.	
DR	PROSITE; PS00222; IGF BINDING.	
DR	PROSITE; PS00484; THYROGLOBULIN 1.	
KW	GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.	
FT	SIGNAL	1 21
FT	CHAIN	22 254
FT	CARBOHYD	125 125
FT	DOMAIN	196 245
FT	SEQUENCE	254 AA; 27745 MW; 3B5316DC CRC32;
QQ		

```

Query Match      4.9%; Score 133; DB 5; Length 254;
Best Local Similarity 39.7%; Pred. No. 5.18e-09;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;

bb      1 mlpfgiivaalllaagrpaplgde-aihhppcseeklarciprvvgceelvrepqgcscatc 59
      | | | | | : | | : | : | : | : | : | : | : | : | : | : | : |
py      6 MGPVRVAFVLLALCSRPAVGQNGSGPCR-CFEPAPRC-PA-GVS-LVID-GCGCCRVC 60

bb      60 aiglmpc 67
      | | | | |
py      61 AKQGLGELC 68

```

RESULT	10	STANDARD;	PRT;	867 AA.
D	SSPO BOVIN			
C	P98167;			
T	01-OCT-1996	(REL. 34, CREATED)		
T	01-OCT-1996	(REL. 34, LAST SEQUENCE UPDATE)		
T	01-OCT-1996	(REL. 34, LAST ANNOTATION UPDATE)		
E	SCO-SPONDI	(FRAGMENT).		
S	BOS TAURUS	(BOVINE).		
C	EUKARYOTA;	ANTIMALIA;	CHORDATA;	VERTEBRATA; MAMMALIA;
C	THELIA;	EUTHERIA;	ARTIODACTYLA;	RUMINANTIA; PECORA; BOVIDAE.
N	[1]			
P	SEQUENCE FROM N.A.			
C	TISSUE=EPENDYMOCYTE;			
X	MDLINE; 96338614.			
A	GOBBON S., MONNERIE H.,			
A	JAMALLE D., DASTUGUE B.,			
L	J. CELL SCI. 109:1053-1061			
C	-(FUNCTION: INVOLVED IN THE MODULATION OF NEURONAL AGGREGATION.			
C	-(TISSUE SPECIFICITY: SUBCOMMISSURAL ORGAN.			
C	-(SUBCELLULAR LOCATION: EXTRACELLULAR.			

CC	-!	DEVELOPMENTAL STAGE: EMBRYO.	
CC	-!	SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.	
CC	-!	SIMILARITY: CONTAINS AT LEAST 2 EGF-LIKE DOMAINS.	
CC	-!	SIMILARITY: CONTAINS AT LEAST 3 LDL-RECEPTOR CLASS A DOMAINS.	
DR	EMBL: X93922; E213357; --		
CC	GLYCOPROTEIN; CELL ADHESION; CALCIUM-BINDING; REPEAT; EGF-LIKE DOMAIN.		
KW	NON TER	1	
FT	DOMAIN	103	EGF-LIKE.
FT	DOMAIN	143	EGF-LIKE.
FT	DOMAIN	180	EGF-LIKE.
FT	DOMAIN	506	LDL-RECEPTOR CLASS A.
FT	DOMAIN	663	LDL-RECEPTOR CLASS A.
FT	DOMAIN	723	LDL-RECEPTOR CLASS A.
FT	CARBOHYD	88	POTENTIAL.
FT	CARBOHYD	309	POTENTIAL.
FT	CARBOHYD	409	POTENTIAL.
FT	NON TER	867	
FT	SEQUENCE	867 AA; 91817 MW; B1224081 CRC32;	
SQ			

	Query Match	4.7%;	Score 128;	DB 9;	Length 867;	
	Best Local Similarity	28.3%;	Pred. 4.95e-08;			
	Matches	32;	Conservative	21;	Mismatches 48;	Indels 12; Gaps 12
Ddb	81 cpdgcpnvtcagelvfhacvpcpltcddisgatoppdr-pcgpp-gwcpagqvlgaq	138	:	:	:	:
Oy	35 CPDEPAPPCPAAGVSLVDGGCCRRV-CAKQIGEL-CT-ERDPCDPHKGLFCDFGSPANRK	91	:	:	:	:
Ddb	139 -grewwpzq-cpclvgdstrywpgqrvtkdqql-cvcqdgrprccpsldcavn	188	:	:	:	:
Oy	92 LGVCT-AKDGA PCIFGTIVRSGESFOSSCKYQCTCLDG-AVGCMF-L-CSMD	140	:	:	:	:

RESULT	11
IDC	IBP5 RAT
AD	STANDARD;
PRT; 271 AA.	
P24594;	
DT	01-MAR-1992 (REL. 21, CREATED)
DT	01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP-5)
DE	(IBP-5) (IGF-BINDING PROTEIN 5).
EN	IGFBP5 OR IGFBP-5.
RATTS NORVEGICUS (RAT).	
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
EUTHERIA; RODENTIA.	
[1]	
SEQUENCE FROM N.A., AND SEQUENCE OF 20-53.	
TISSUE=OVARY;	
MEDLINE; 9124847.	
SHIMASAKI S., SHIMONAKA M., ZHANG H.-P., LING N.;	
J. BIOL. CHEM. 266:10646-10653(1991).	
[2]	
SEQUENCE FROM N.A.	
STRAIN=SPRAGUE-DAWLEY;	
MEDLINE; 93176146.	
ZHU X., LING N., SHIMASAKI S.;	
BIOCHEM. BIOPHYS. RES. COMMUN. 190:1045-1052(1993).	
-1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs	
AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH	
PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE	
INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.	
-1- TISSUE SPECIFICITY: MOSTLY IN KIDNEY.	

-!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.

CC EMBL; M62781; G204746; -.
 DR EMBL; L08275; E73333; -.
 DR PIR; A40403; A40403.
 DR PIR; JCI1463; JCI1463.
 DR PIR; F40403; F40403.
 DR PROSITE; PS00222; IGF BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN_1.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 271 INSULIN-LIKE GROWTH FACTOR BINDING
 FT CHAIN 20 271 PROTEIN 5.
 FT DOMAIN 214 262 THYROGLOBULIN TYPE 1.
 FT SEQUENCE 271 AA; 30298 MW; 0AA79506 CRC32;
 Query Match 4.6%; Score 125; DB 5; Length 271;
 Best Local Similarity 36.0%; Pred. No. 1.88e-07;
 Matches 27; Conservative 12; Mismatches 28; Indels 8; Gaps 7;

Db 7 llllaacavpaqlgsfvhcepdckalemcpsplgceivkepgcgcmcalaeagasc 66

Qy 14 WLLALCSRPVAVQNCSPGCR-PCPEPAPRC-PA--GVSLVLD-GCGGCRVCARQLGELC 68

Db 67 gvytete-ag-glrc 79

Qy 69 TE-RDPCDPKGLFC 82

RESULT 12
 ID IBP4 HUMAN STANDARD; PRT; 258 AA.
 AC P22682;
 DT 01-AUG-1991 (REL. 19, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
 DE (IBP-4) (IGF-BINDING PROTEIN 4).
 GN IGFBP4 OR IBP4.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91186988.
 RA LATOUR D., MOHAN S., LINKHART T.A., BAYLINK D.J., STRONG D.D.;
 RL MOL. ENDOCRINOL. 4:1806-1814 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE; 91133415.
 RA SHIMASAKI S., UCHIYAMA F., SHIMONAKA M., LING N.;
 RL MOL. ENDOCRINOL. 4:1451-1458 (1990).
 RN [3]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-41.
 RC TISSUE=OSTEOSARCOMA;
 RX MEDLINE; 91225006.
 RA KIEFER M.C., MASTIARZ F.R., BAUER D.M., ZAPF J.;
 RL J. BIOL. CHEM. 266:9043-9049 (1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;

RA STRONG D.D., MORALES S., LEE K., BOONYARATANAKORUKIT V.,
 RA BAYLINK D.J., MOHAN S.;
 RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [5]
 RP SEQUENCE OF 22-53.
 RC TISSUE=COLON;
 RX MEDLINE; 91235178.
 RA CULOUSCOU J.-M., SHOYAB M.;
 RL CANCER RES. 51:2813-2819 (1991).
 CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -!- BINDS IGF-II MORE THAN IGF-I.
 CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
 DR EMBL; M38177; -; NOT ANNOTATED_CDS.
 DR EMBL; M62403; G184816; -.
 DR EMBL; U20982; G695254; -.
 DR PIR; A36549; A36549.
 DR PIR; B37232; B37232.
 DR PIR; B39842; B39842.
 DR MIM; 146733; -.
 DR PROSITE; PS00222; IGF BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN_1.
 KW GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 21
 FT CHAIN 22 258 INSULIN-LIKE GROWTH FACTOR BINDING
 FT CHAIN 22 258 PROTEIN 4.
 FT CARBOHYD 125 125 POTENTIAL.
 FT DOMAIN 200 249 THYROGLOBULIN TYPE 1.
 FT CONFLICT 51 51 P -> A (IN REF. 1, 4 AND 5).
 FT CONFLICT 198 198 I -> F (IN REF. 1 AND 4).
 SQ SEQUENCE 258 AA; 27934 MW; 58AC8AC3 CRC32;

Query Match 4.5%; Score 123; DB 5; Length 258;
 Best Local Similarity 38.2%; Pred. No. 4.54e-07;
 Matches 26; Conservative 14; Mismatches 22; Indels 6; Gaps 6;

Db 1 mlpclvaalllaagpgpslgde-ahhpcpcseeklarcpvpgceelvrepqgcacac 59
 Qy 6 MGPVAVFVLLALCSRPVAVQNCSPGCR-CPDEPAPRC-PA-GVS-LVLD-GCGGCRVC 60
 Db 60 alglgmpe 67
 Qy 61 AKQLGELC 68

RESULT 13
 ID 60A DROME STANDARD; PRT; 455 AA.
 AC P27091;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE 60A PROTEIN PRECURSOR.
 GN 60A OR IGFB-60A.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92021021.

RA	WHARTON K.A., THOMSEN G.H., CELBART W.M.;
RJ	PROC. NATL. ACAD. SCI. U.S.A. 88:9214-9218(1991).
RL	[2]
RN	SEQUENCE FROM N.A.
RP	MEOLINE; 92290120.
RX	OOCOTR J.S., JACKSON P.O., RASHKA K.E., VISALLI M., HOFFMANN F.M.;
RA	DEV. BIOL. 151:491-505(1992).
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH PEAKS
CC	OF TRANSCRIPTION DURING EARLY EMBRYOGENESIS, IN POPAE, ANO IN
CC	ADULT MALES.
CC	-1- SUBUNIT: HOMODIMER, OISULFIOE-LINKED.
CC	-!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
OR	EMBL; M7012; G156728; -.
OR	EMBL; M84795; G156730; -.
DR	PIR; A41233; A41233.
DR	PIR; A43918; A43918.
DR	HSSP; P08112; 1TFG.
DR	FYLBASE; FBGN004788; TGF-BETA-60A.
DR	PROSITE; PS00250; TGF BETA.
KW	GROWTH FACTOR, CYTOKINE; GLYCOPROTEIN; SIGNAL.
FT	SIGNAL ?
FT	PROPEP ?
FT	CHAIN 336
FT	OISULFIO 354
FT	OISULFIO 420
FT	OISULFIO 383
FT	OISULFIO 387
FT	DISULFIO 419
FT	CARBOHYD 238
FT	CARBOHYD 250
FT	CARBOHYD 396
SQ	SEQUENCE 455 AA; 51687 MW; 6FAAC7C0 CRC32;
Ob	Query Match 4.5%; Score 121; OB 1; Length 455;
Qy	Best Local Similarity 43.2%; Pred. No. 1.09e-06;
	Matches 16; Conservative 9; Mismatches 10; Indels 2; Gaps
	419 ccaprtgialpvhlndenvnlkkyrnmivkcgcgh 455 : : : : : : : : : 292 CCTHRITTLVEFKCPDGEV-MKKQM-MFIKTCACh 326
RESULT 14	
IIO IBP4 MOUSE STANDARD; PRT; 254 AA.	
AAC P47879;	
DDT 01-FEB-1996 (REL. 33, CREATEO)	
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)	
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)	
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4) (IBP-4) (IGF-BINDING PROTEIN 4).	
GG IGFBP4 OR IGFBP-4	
N GN MUS MUSCULUS (MOUSE).	
OS MUS MUSCULUS (MOUSE).	
CCC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
CCC EUTHERIA; ROENTZIA.	
RN [1]	
RN SEQUENCE FROM N.A.	
RRC TISSUE=LIVER;	
RRC MEDLINE; 95121750.	
RX SCHULLER A.G.P., GROFFEN C., VAN NECK J.W., ZWARHOFTH E.C.,	
RX OROP S.L.S.;	
RA MOL. CELL. ENDOCRINOL. 104:57-66(1994).	

1	mlpcslvaalltagrrpalsgde-aihpccpccsekclarcrrpvcceelvrepccgcsatc	59
6	MGPVRFVAFVLLALCRPAVGQNCSPCR-CFOEPAPRC-PA-GVS-LVLD-GCGCCRCVC	60
60	algiq 64	
61	AKQLG 65	
Query Match	4.3%;	Score 116; OB 5; Length 254;
Best Local Similarity	36.9%;	Prd. No. 9.35e-06;
Matches	24;	Conservative 14; Mismatches 21; Indels 6; Gaps 6;
Ob	1	mlpcslvaalltagrrpalsgde-aihpccpccsekclarcrrpvcceelvrepccgcsatc
Qy	6	MGPVRFVAFVLLALCRPAVGQNCSPCR-CFOEPAPRC-PA-GVS-LVLD-GCGCCRCVC
Ob	60	algiq 64
Qy	61	AKQLG 65
RESULT	15	
IO	IBP1 RAT	STANOARD; PRT; 272 AA.
AD	P21743;	
AC	01-MAY-1991 (REL. 18, CREATEO)	
TOT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)	
OT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)	
OE	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (IGFBP-1)	
OG	(IBP-1) (IGF-BINDING PROTEIN 1).	
OE	IGFBP1 OR IGFBP-1.	
OS	RATTUS NORVEGICUS (RAT).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC	EUTHERIA; RODENTIA.	
[1]		
RN	SEQUENCE FROM N.A.	
RP	TISSUE=OECIODA;	
RC	MEOLINE; 90231347.	
RR	MURPHY L.J.; SENEVIRATNE C.; BALLEJO G.; CROZE F.; KENNEDY T.G.;	
RL	MOL. ENDOCRINOL. 4:329-336(1990).	
[2]		
RN	SEQUENCE FROM N.A.	
RP	TISSUE=LIVER;	
RC		

RX MEDLINE; 91141487.
RA MOHN K.L., MELBY A.E., TEWARI D.S., LAZ T.M., TAUB R.A.;
RL MOL. CELL. BIOL. 11:1393-1401(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93149132.
RA OOI G.T., TSENG L.Y.H., TRAN M.Q., RECHLER M.M.;
RL MOL. ENDOCRINOL. 6:2219-2228(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE; 94250701.
RA LACSON R.G., OEHLER D., YANG E., GOSWAMI R., UNTERMAN T.G.;
RL BIOCHIM. BIOPHYS. ACTA 1218:95-98(1994).
RN [5]
RP SEQUENCE OF 26-59.
RX MEDLINE; 90322923.
RA UNTERMAN T.G., OEHLER D.T., GOTWAY M.B., MORRIS P.W.;
RL ENDOCRINOLOGY 127:789-797(1990).
CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -1- BINDS EQUALLY WELL IGF-I AND IGF-II.
CC -1- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR EMBL; M58634; G204733; -.
DR EMBL; M89791; G204737; -.
DR EMBL; L22979; G1098473; -.
DR FIR; A36082; A36082.
DR FIR; A37398; A37398.
DR FIR; A39683; A39683.
DR PROSITE; PS00222; IGF BINDING.
DR PROSITE; PS00484; THYROGLOBULIN_1.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 23
FT CHAIN 26 272
FT INSULIN-LIKE GROWTH FACTOR BINDING
FT PROTEIN 1.
FT DOMAIN 215 264
FT SITE 259 261
FT CELL ATTACHMENT SITE.
FT CONFLICT 79 79
FT R -> A (IN REF. 1 AND 4).
FT CONFLICT 111 112
FT AA -> PP (IN REF. 1).
FT CONFLICT 201 201
FT A -> R (IN REF. 3).
FT CONFLICT 265 265
FT H -> N (IN REF. 1).
SQ SEQUENCE 272 AA; 29684 MW; 59EDF790 CRC32;

Query Match 4.2%; Score 113; DB 5; Length 272;
Best Local Similarity 42.9%; Pred. No. 3.32e-05;
Matches 21; Conservative 4; Mismatches 18; Indels 6; Gaps 6;

D0 43 cpvpaas-cpe-lsrp-acgcgcptcalplgaacgvatarc-aq-glsc 86
QY 35 CPDEFAPRCFAGVSLVLDGCGCCRCVCAKGLGELC-TERDPCDPHKLFC 82

Search completed: Wed Sep 17 09:28:33 1997
Job time : 51 secs.